

Sample of current work toward a book/thesis on modality
Tristan Haze

Contents (will not correspond to the book)

Page numbers are approximate

1. Metaphysical modality: the necessary aposteriori and the contingent apriori.....	p2
2. The problem of reference/representation.....	p28
3. Epistemic modals.....	p44
4. Apriori-aposteriori.....	p62
5. Miscellaneous.....	p64
6. On method.....	p74

Metaphysical Modality

"If there were a maximally good conceptual scheme, it would sustain by itself all necessities, and nothing besides". (Some of our problems with modality resemble more familiar problems with truth - with Peircean pragmatism, etc.)

The 'truth wars' - subjectivists/relativists vs. objectivists/absolutists - are paralleled by 'modal wars' - skeptics, conventionalists and deflators vs. substantialist metaphysicians.

What is the status of the claim with which Lewis begins his book? (That there are other ways the world could have been other than the way things are.) We have a tendency to hear it in the wrong key. Imagine someone who had lost a loved one saying 'Things really could have been different'. (There seems something self-deluding in this, although it is no ordinary self-delusion. You don't fix it with negation.)

World -> concepts -> necessity.

But what of these arrows? The world doesn't somehow rationally force particular schemes on us, and necessities are no mere epiphenomena of schemes. One thing is clear: a Lewisian "concept-free" view and an Ayerian "world-free" view are both wrongheaded. I have leant toward the latter myself, but I now want to get beyond mere partisanship and to take another road entirely.

Cutting through rationalism vs. empiricism, substantialism vs. conventionalism, about modality. My view: modality is bound up with concepts, but 'right' concepts. And their status as such cannot be called 'mind-dependent'. Right concepts - but to say *the* right concepts borders on dogmatism, on stupidity.

* * *

Ayer wanted to say: the necessities fall out of our linguistic/conceptual scheme. Codify 'the rules of correct English' (*LTL*, find page REF) and you've codified the necessities (or a base from which they may be logically derived). But the Kripkean necessary a posteriori, semantic externalism etc. make it clear that one needs a lot more fine structure than 'rules of correct English', and that some structures are in some sense not correct, in a way which has nothing to do with internal incoherence, but rather to do with inadequacy to an external referent.

Thus, my view of modality retains part of Ayer's picture: the (metaphysical) necessities are seen as things which in some sense fall out of conceptual systems - things which are invariant through all configurations of the system - roughly in the manner of the tautologies of logic. But not just any system will do. The elements of the system must be adequate to their objects.

Our conceptual system has different configurations, has moving parts. Here there is a queer tendency to say something like: yes, but this is only a practical expedient which we adopt because we fall short of knowing everything.

'All truths are necessary - the mind of God needs no moving parts.' - OK, but does that mean that there is no important difference between what we call contingent and what we call necessary? Between 'Water is H₂O' and 'There is some water in this cup'? Between ' $2 + 2 = 4$ ' and 'Here it is written that $2 + 2 = 4$ '? That doesn't seem to be something the necessitarian would want to admit to, in so many words. It seems rather

more like the point is that none of this would quite correspond to anything in God's mind. The necessitarian thinks, as it were: OK, it doesn't correspond, but what *is* in God's mind is more like ' $2 + 2 = 4$ ' in our mind than 'Here it is written that $2 + 2 = 4$ ' in our mind.

Here I think we are able to see the point of Wittgenstein's remark in *RFM* about concepts in necessary propositions having to have a meaning in non-necessary ones. I.e. If God's concepts are perfectly static and rigid, then in what sense are they concepts at all?

This is what is so strange about the thought that our conceptual systems only have moving parts because we don't know everything. Isn't it essential to a conceptual scheme that it has moving parts? Is our thought, then, just that we only have conceptual systems at all because we don't know everything? But how can one know *anything* without a conceptual system? (Omniscience as an idea that 'can't be worked out'. Psycho-panteism in this connection: God = mind = world.)

'Everything happens for a reason'.

Liebniz had quite a task in avoiding Spinozistic necessitarianism, as can be clearly seen in the *Monadology*. Compare the two thinkers on the problem of evil: for Spinoza, there's - in a sense - really no such thing: it's perspectival. In another sense, however, all is good. According to Liebniz on the other hand, we of course have the best of all possible worlds. Curious: in some sense, it seems like given *that*, some sort of Spinozistic state of affairs holds contingently - or more accurately, by the grace of God. Actual evil - for Liebniz - is *in a sense* not substantial evil: for nothing better is possible. Inferior possible worlds, on the other hand, would have contained evil of a more substantial - because possibly-absent - kind.

Liebniz's problem here is, at root, similar in form to the problem of whether God could create a stone he couldn't lift: if it is of the essence of God that he is maximally good, it would seem that if he created a world at all, it must necessarily be the best of possible worlds. But if that's necessary, how are the others *possible* worlds? (Is the solution here to say that God might not have existed? Or might not have had His essence?)

It is a bit strange that I am even talking about this.

* * *

'I went to X, but could have gone to Y instead.'

How might such a proposition be used? A retrospective epistemic use. Or to point to various facts - e.g. that Y wasn't too far away, or that I was under no compulsion to go to X, or that I was seriously considering going to Y, etc. But what if I say that none of that is quite to the point? 'Sure, those things might be true, but over and above that, I'm saying that I really could have gone to Y. The world could have gone differently here, and my going to X was just how it happened to turn out.' What could this mean? (We can imagine someone being bothered, preoccupied, by the thought that things really could have gone differently.)

(1) There are things which could have gone otherwise.

(2) There are contingencies.

When a philosopher asserts (1) or words to that effect, typically the reaction they want is 'of course!'. Now, this would happen sometimes, but more common among non-philosophers might be perplexity - 'what do you mean?' - and skepticism - 'but how do you know?'. (I have some experience of this.)

Both in and out of serious philosophy, there is a tendency to take (1) and (2) in what I would call the wrong way. Roughly: as though they described something extremely fundamental and general about reality. And

yet that is not an entirely wrong-headed thought, for it is immensely important that we talk and think about non-actual scenarios.

When one hears (1) and (2) in an "inflated" way - whatever that comes to - necessitarianism can seem like the antidote: something wrong has been said, and so it should be denied. But then the denial turns out to be just as troublesome, if not more! (Now it is trivially false, instead of trivially true - like Idealism in relation to Realism.)

The disease is not fully containable in any one proposition.

It is as though, when someone pushes a statement of counterfactual possibility in a certain way, we take them to be making a very strong claim, quite beyond our ken.

It can be very important to realize that some propositions are necessary - 'One cannot feel another's pain', 'The opium worked (if it worked at all) by means of a dormitative power.'. Likewise, contingent - 'There are two sexes', 'We have one body each', 'We use the decimal system to count.' It can be very instructive to get clear about the "modal status" of these things. And of course this does not mean leaving the view that they are contingent (or necessary) and coming over to the other side, but rather: noticing something, which one might have been quite blind to. Think of the conceptual development of children - one cannot truly say that they go through a phase of believing that everyone knows the same as them, but we can see what this is getting at. In the place where we have a fundamental distinction or concept of great importance in adult life, they do not have anything to speak of.

So: it can be important to realize, as we say, the modal status of certain propositions. But - and this is crucial - it does not *at all* follow from this that it is important, desirable or possible to draw a line which clearly divides all propositions according to modal status.

What is the modal status of 'Our conceptual scheme has moving parts'? I.e. how could anything else be a conceptual scheme? (Mind of God.) And yet the proposition easily acquires the flavour of a general fact of nature. Why? Well, lots of important facts lie immediately underneath.

* * *

It is of fundamental importance to this investigation that in certain situations it makes sense to talk about "having the right concept", or the rightness of a given concept.

* * *

We philosophize about truth and modality and think something like: truth is what we should ideally believe, modality is how we should ideally conceive. Or again: the truth is the position one should ideally put one's model of the world into, and the range of possibilities is the range of movement, or the structure, which one should ideally give to the model.

This has its value, but *then* leads to problems about the meaning or source of the ideal.

Here it is of the utmost importance to note that whatever our initial thought amounted to, it was *not* that 'x is true' *means* 'x is believed ideally' - for that form is already used up, to mean something else. (This device - 'already used up' - I got from a transcript of a lecture Wittgenstein delivered in the early 1930s.)

((There seems to be a connection between this and what is going on with the contingent *apriori*: '1m' doesn't mean 'the length of S at t', and yet there is a conceptual connection - something more than just a coincidence of reference. - 'is true' can never mean 'is believed ideally', and it's wrong to say they just happen to coincide, but there is a middle way - where, for instance, we see ourselves as regulating the pre-existing sense of 'is believed ideally'. In both cases, we end up with *more* than mere extensional coincidence, but still not *identity* of sense. There are, however, significant differences between the case of '1m'/the length of S at t' and that of 'is true'/is believed ideally. Difference: '1m' was defined for the first time. To assimilate the two a little: think of regulating an inexact unit - '1 pace'.))

(What is said for 'is true'/'is believed ideally' above also applies to 'is metaphysically necessary'/'is invariant through configurations of a relevantly adequate conceptual scheme'.)

* * *

'According to current orthodoxy, metaphysical possibility can neither be reduced to, nor eliminated in favour of, linguistic rules and conventions; it constitutes a fundamental, mind-independent matter for thought and talk.'

- vii, Introduction to volume on Conceivability and Possibility

Two ways of interpreting such rhetoric about modality:

'The conventionalism had to go.' - right.

'Consideration of language and concepts is not to the point.' - wrong.

* * *

The notion of the mind of God. It can look as though everything is an approach to this ideal, and yet we have absolutely no grip on the ideal. We (i.e. myself and most active philosophers) do not have any positive belief in the mind of God (in the old sense), and all the minds we do believe in are completely un-Godly.

Spinoza's view that all truths are necessary.

Our conceptual scheme has moving parts. But if God knows all, he doesn't need any moving parts.

So now it looks like, in a sense, the ideal conceptual scheme has no moving parts - no need to dance further with experience or reason. And yet we will always need a conceptual scheme with moving parts, and can really make no sense of anything else. Thus how could the mind of God be *our* ideal?!

From this perspective, one could imagine pantheism as a kind of scholastic solution to our present problem. (For it would block the assimilation of our mind with "God's".) ('A full, non-redundant model of an object must be a duplicate of the object, or the object itself.')

We say things could have been otherwise - I could have walked somewhere else today. But then someone says 'but could that *really* have happened?' and it's like the stakes are raised somehow.

"What makes you think things could have *really* been different, and that this isn't just a function of our ignorance?" - The first thing to say is there's no simple answer here: one can't just point to evidence. The question is symptomatic of something.

* * *

Notes on Carrie Jenkins' recent work:

p 16 - 17

Here Jenkins introduces talk of rightness of concepts, fitting, etc.

"Grounded concepts can be thought of as a trustworthy map of the world: they represent the world as it really is, and it is no accident that they do so. Therefore, I think, we can recover information about the world by examining this map. Those pieces of information that we can so recover are the ones which I want to describe as 'following epistemically' from the concepts in question."

When Jenkins speaks of concepts fitting things, I'm happy. But then what she does with this down the track worries me. She talks of a conceptual scheme as being like a picture, and talks about concepts lining up with the individuals and properties in the world. Here I'm reminded of Quine's phrase 'museum myth'. Perhaps another way of registering this is to say: it looks like perhaps Jenkins is here saying (dogmatically, to my mind) that there is ultimately just one true way of conceiving. That's probably not quite fair, but still,

it can sort of look that way.

Just because there are x's (is F'ness) does not mean that every mind should have a concept which could be described as the concept of x's (F'ness). When we say there are x's, we are using our scheme.

Here we must of course resist the urge, which doesn't seem to go away even when one has realized this for years, to think of the proposition 'there are x's' as a proposition *about* the scheme. For then, to be really consistent, we ought to interpret propositions about the scheme as being propositions about how the scheme should represent the scheme.

Why do we wish to make this transposition? We reflect that radically different conceptual schemes might be possible, then we reflect on our existence-assertions, and feel a desire to as it were make room for our first reflection in these assertions. We now feel the need to say something which would be neutral between our scheme and other, radically different possibilities. And thus we think of saying something like: 'To say "there are x's" is to say that, given the general orientation of the present framework (and assuming, perhaps, that the whole thing isn't hopeless) it ought to have a concept of x's.'

The first thing to note here is that this shift is clearly uncalled-for in many (or most, or all) practical cases. But perhaps there are cases where we would want to say the second, qualified thing, when we *wouldn't* want to make the corresponding existence-assertion: for example when discussing a theoretical framework which may turn out to be quite inadequate as a whole - in biology for example.

* * *

Sidelle (1989, 1-2) describes his view as one on which "modality does not find its home in the mind-independent world, but rather in us, in our ways of speaking and thinking".

Also, Wittgenstein's 'we predicate of the thing that which lies in our mode of representing it'.

On the other hand, we have ((quoted above also))

'According to current orthodoxy, metaphysical possibility can neither be reduced to, nor eliminated in favour of, linguistic rules and conventions; it constitutes a fundamental, mind-independent matter for thought and talk.'

- vii, Introduction to volume on Conceivability and Possibility:

It is clear that siding with either of these ideologies will cripple our thinking.

* * *

I can imagine someone objecting to some of the things I've been writing about modality along these lines: "You talk of concepts and conceptual schemes and the like, and their adequacy or approach to some queer ideal, as if to somehow ground modality in that. But how is that any more attractive than Peirceanism about truth - the view that what is true is what is believed at the ideal limit of inquiry?"

It would be very easy for me to go wrong in replying to this.

((This belongs with the more recent stuff about being able to say more about modal character, when the question of truth - adequacy - is put to one side.))

* * *

Liebniz, Monadology, sect. 43:

It is farther true that in God there is not only the source of existences but also that of essences, in so far as they are real, that is to say, the source of what is real in the possible. For the understanding of God is the region of eternal truths or of the ideas on which they depend, and without Him there would be nothing real in the possibilities of things, and not only would there be nothing in existence, but nothing would even be possible. (Theod. 20.)

We are pushed into thinking about 'the understanding of God', and yet the whole role of our understanding - the only understanding we know - is "constitutively tied" to the fact that we are trying to get on in the world, are not gods.

* * *

How could the real possibilities of things come down to concepts - and if they do in some sense, isn't this a kind of lucky accident? Could there not be things which are in principle not conceptualizable?

I must avoid positing any pre-established harmony between all the possible concepts and all the possible realities.

Here, a sort of Wittgensteinian thing to say would be: there is a kind of harmony, but made in language, by grammar. (REF - Zettel, I think? begins something like 'As with everything metaphysical...'.) (I'm not so sure. That seems right in a way, but the grammatical connection, or connections, in question are probably more complicated than it might look at first. And also: why is grammar that way?)

In dreams it can happen that we are with a certain person whom we know very well, and we know it is them, even though they take a quite different bodily form from in reality. One doesn't realize the oddness of this in the dream, but upon thinking of it later may suddenly be struck by the strange fact.

It is noteworthy that one can even *imagine* some kind of queer dream experience in which an animal or even a non-living thing seems to *be* some particular person one knows. Or in which a thing seems to *contain* someone's spirit.

How odd that the phrase 'dreaming that one's family were grandfather clocks' describes something which might happen to someone! Is it not extraordinary that that this absurd conception (i.e. of one's family being grandfather clocks) might actually be quite *essential* for the purposes of describing and understanding a reality? (For the dream, qua dream, would be a reality.)

(This is connected with Wittgenstein's remarks about how it can look as though such reports about the contents of thought and experience are somehow indirect - that thought and experience is here being described by means of inessential connections it happens to have with other things, and that a more direct description ought to be possible.) (REF Phil Grammar?)

If it makes sense to dream that *p*, '*p*' cannot be nonsense. However, it looks like it can very easily be metaphysically impossible.

This has implications for what we say about the way dream reports are represented. It looks as if we make use of a conceptual system which is not to be endorsed, rather than our endorsed system (whose configurations we endorse as metaphysical possibilities). Or perhaps better: it looks as if we are loosening up our conceptual system, lifting certain rules and restrictions. However, we seem to need this unendorsed or loosened system in order to correctly represent what was really dreamt. Thus it should be given its proper recognition as a fully legitimate part of our representational activity, not some second-rate thing.

After all, it seems it is metaphysically possible to dream the metaphysically impossible. So what looked like our "primary" system, the one which we construct and restrict so as to give us just the metaphysical possibilities, does not really give us all of them, since it cannot by itself represent the metaphysically impossible dream content of metaphysically possible dreams. To represent such dreams, it seems like this "primary" system needs to work in tandem with a looser version of itself (or an alternate system).

(A less fantastic, but very similar point, comes up with things like 'John believes that Hesperus is not Phosphorus'. The similarity is that we need something from outside our "primary system", the difference is that there is nothing outlandish or loose about the alternative; it is just something we think isn't true.)

((NBB There is a spot somewhere where I raise a similar problem, but go for a more metalinguistic understanding of it. I should probably connect the two.))

In a similar connection: couldn't some forms of understanding be essentially parasitic on various misunderstandings? I.e. could there be things which could not be understood by an ideally rational agent, although understandable by lesser agents? And what would an ideally rational agent be here? One which strikes the best compromise so as to understand as much as possible, or one which avoids all misunderstanding?

Here perhaps we get a glimpse of how understanding is bound up with life - how it is not the hovering, detachable, ethereal thing we are inclined to regard it as.

('You might be able to *simulate* the thoughts of the Greeks quite well after long study, but since you are a modern, it will not be the same - you will never truly understand their thoughts as they understood them.' - sometimes 'understand' is close to 'have', other times far from it.)

What would it be like to have perfect, or even very great, understanding for one second? Isn't this nonsense? 'No, it makes sense - it holds if, had the state gone longer, and given the proper surroundings, great understanding *would have* been manifested externally.' - But what is 'the state' here, and what is the criterion for it going longer, as opposed to giving way to another state?

* * *

My investigation so far has led me to want to say something like: the (metaphysically) necessary truths are those which are fixed by the concepts involved in them, where the concepts are adequate to their objects. Or again: if you're given adequate concepts of things, you thereby get the necessary truths about those things. Or again: your concepts yield necessary truths insofar as they are adequate to reality.

'Adequate' here doesn't mean something like 'good enough for practical purposes', because something false (or falsity-inducing) could fulfil that requirement. What does 'adequate' mean here, then? Well, it was initially introduced to deal with the Kripkean necessary a posteriori. Consider the case of Hesperus and Phosphorus. Someone who thinks that Hesperus is not Phosphorus has a conceptual system with two separate individual concepts. Such a conceptual arrangement is in itself perfectly alright, but given that both concepts are in fact *of* the same things (and this, following Putnam with Twin Earth, is not something determined internally by the system), it is inadequate. A conceptual system is inadequate with respect to some matter, we might say, if it cannot yield the truth about that matter. We distinguish 'mere false belief' from false belief which involves conceptual inadequacy. And false belief in identities and negated identities involving proper names (individual concepts) is never mere.

At least, it is helpful to carve things up this way. I am not suggesting that it is the only way. For example, all of this will sound wrong if one thinks of conceptual adequacy as something which can be rationally checked from inside. 'Conceptual adequacy' is here being used in a way which goes beyond a priori coherence. Another example: one can individuate conceptual systems more coarsely, so that in changing from thinking that Hesperus is not Phosphorus to thinking that it is, one doesn't change one's system, but only the configuration of one's system. The above idea of adequacy involves a finer-grained conception.

Secondly, all of this talk has to be taken in the right spirit. It is not trying to explain, ground, justify or account for things, so much as to describe them in a helpful way.

If we falsely believe in two things where there is just one, we might have two concepts which, internally speaking, would be adequate if the outside world were different. But if we're agnostic as to their identity, then our conceptual situation is, so to speak, internally incomplete - it would not be fully adequate in any surrounds. Similarly, in the case of kinds and their constitution: if we falsely believe water is XYZ, our internal situation might be adequate in different surrounds (Twin Earth), whereas if we're agnostic as to what constitutes water, we are in an internally incomplete conceptual situation.

* * *

'If (metaphysical, subjunctive) necessity comes to being fixed by concepts adequate to their objects, then to get all the necessities, what you need is an adequate concept for all the things there are (in the broadest sense of 'thing' - i.e., not merely what we would call 'individuals' but things like addition, communication, etc.).'

I think something has gone wrong, or is going wrong, here.

The notion of alternative conceptual schemes. An ambiguity in the notion of 'all the truths' - scheme-relative vs. all schemes.

* * *

My talk about the adequacy of concepts to their objects does not need to be restricted in any way, although it obviously has a potential to be misleading in, say, mathematical cases, particularly when the object is a mathematical 'individual' such as a number.

It really comes down to 'adequacy, given the situation both inside and out'. To be sure, this talk doesn't go very far, but it goes far enough to be worthwhile. (There is a tendency to try to give it a kind of role it cannot play.)

* * *

"What are the ways things could have been?" One pernicious kind of confusion when talking about metaphysical necessity - the modality identified, and separated from other concepts, by Kripke - is to think of it in a manner which would be appropriate for some kind of deep and unimprovable-upon form of epistemic possibility. This is illustrated by the notion of the Creation as a kind of event. God did something, and it may have been reasonable to expect this to go various ways. It is important for our concerns here that such an idea is *natural* to some degree.

We get the feeling that any kind of conceptual view of necessity, even if we grant everything to the concepts including external adequacy, somehow makes modality flimsy, shadowy - somehow still overlooks reality in some way. "The way things *really* could have been, concepts aside." (The notion of an inconceivable object. If such a thing is impossible, how is that not just a piece of luck? It's as though we are presupposing some kind of pre-established harmony between possible concepts and possible realities, and of course this would be highly suspicious.) NBB

* * *

The 'metaphysically necessary' is marked out by its logical character, not by subject matter. Thus phrases like 'knowledge of metaphysical necessity' - and the question of 'how we have such knowledge', etc. - are very apt to confuse.

There is a strong tendency in philosophy to think of knowledge of how the world *must* be as penetrating deeper into the world than mere contingent knowledge. As though we saw how a machine behaved, but then looked into its workings and saw that it had to behave that way. Or, we form a hypothesis about the workings based on, and in order to make sense of, the behaviour. (This gets us into a lot of difficulties, but it would be stupid to call it incorrect - or correct for that matter.)

I have been talking about the relation of metaphysical modality to representation and the theory of reference, as well as its relation to epistemic modality. In this, leading themes have been concepts, individual concepts, adequacy of concepts, structure vs. configuration of structure, and leaving things open vs. holding them fixed. But the purpose, or place, of all this discussion may not be clear. It is not material for an ontological grounding. (That path would lead me to the mind of God, or something similar.) Neither is it material for a definition or a conceptual analysis.

I can characterize (figuratively) the difference between propositions of metaphysically necessary character and those of metaphysically contingent character, but I can't say anything "substantial" about what makes a proposition with necessary character *true*. All I have there is the notion of adequacy of concept to object - and I feel it is important that I do not try for more. This notion of adequacy is not, as it were, a portal to a new conceptual world. It lies very close to the notions of truth and metaphysical necessity, and its value in

our considerations is that it helps us give the kind of description which will clarify things. We could say: its value lies in its logical form, rather than its matter.

Thus, while it might look as though the way forward now would be to give some kind of independent account of adequacy of concepts, this is not actually so - although of course there *are* things we can say to clarify the notion.

* * *

Suppose we say that what characterizes metaphysically necessary propositions is that we take a certain attitude toward them, regarding them as true while in a sense excluding their negations from our system. There seems to be something in this, but it seems inadequate as well. For it is not as though the proposition sits there by itself, without modal status so to speak, until we take some attitude to it. This is wrong obviously for the trivial reason that it makes modal status a temporal matter, but also because it fails to do justice to the fact that the 'modal status' is part of the essence of the proposition.

Here, by 'the proposition' I don't mean some object totally abstracted from any symbolic system, nor just the sign, but rather the sign in use. Although perhaps I can't make this distinction in quite the Tractarian way: the sign vs. the sign in its projective relation to the world. I can't do this, one might say, because the whole point about necessary propositions is that in a sense there *is* no projection. Within a conceptual system, a necessity holds in all configurations, so it cannot - at the level of configurations, so to speak - correspond or fail to correspond with anything. In this way, necessities and impossibilities are analogous to Wittgenstein's categories of tautology and contradiction.

What happens, though, to the stuff about tautologousness or contradictoriness being shown by the symbol alone? Transcribing this into modal terms (and putting aside any considerations of effective decidability in logic) we get - on my view - a half-truth. If you know and fix on a particular system, you can determine the modal character of a proposition without inspecting the world. However, you cannot thus tell that your system is adequate to the world. So you may posit two systems, for example, and then look at the world to choose between them. In this way, it might be said that you get a kind of higher-order projective relation, holding between the scheme - or parts of the scheme - and the world, as opposed to one holding between particular configurations of the scheme and the world. And it would be this kind of projection which is in question with empirical identity statements.

Thus we *can* speak of correspondence here, but at the level of structure rather than configuration. (At the level of configuration, a believed empirical identity ' $a = b$ ' is fixed - a could not have been distinct from b . However, one may consider whether a change in structure is called-for - whether a might not actually be distinct from b after all.)

The above is expressed in a way which might arouse suspicion. On this, I want to say: the difference between the "two orders" - structural adequacy and adequacy of a structure-configuration - is not so great as my pictures and metaphors might suggest. There is, for example, no reason why we might not also conceptualize what I call 'different structures' as different configurations, at a higher level, of some superstructure, some more coarsely individuated structure. But I express things as I do, in order to put a spotlight on the distinction between metaphysical necessity and contingency.)

* * *

With our conceptual systems we do such and such, and we do it rightly iff ...

To fill the blank after the 'iff', Leibniz and Spinoza used the notion of the mind of God.

But I don't need this second conjunct at all. I just want to say: we do such and such.

As for whether it is right or not - in general, what does this mean?

I simply have nothing to say which is of that kind. I don't believe there is anything to say.

* * *

We must fight the urge to see our modal talk as more "grounded" than it is. (What does that mean? Why would we do that? - Well, we might try to see it as grounded in a way analogous to certain parts of

non-modal talk, and this might cause trouble.)

'X is possible.'

For many substitutions, we're not sure what to say, or what could decide the matter. Thus arises a feeling that the 'outer reaches' of possibility are a realm which we have scant knowledge of. What is *really* possible, irrespective of what we think.

An angle on this: sometimes, we are faced with various conceptual alternatives, and while it may be that one of them is definitely superior to the others, we are unable to determine a clear winner, perhaps due to limitations of cognition, or of experience. (I am not saying there aren't cases where there really isn't a clear winner, but simply choosing to focus on cases where there actually is, although we don't know it.) It is not hard to see how such a situation could make the answers to certain possibility-questions seem out of reach.

(We perhaps get something like this situation in the philosophy of time travel. But we probably also get the other sort of situation, which I ruled out in the previous parenthesis. E.g., maybe with 4D-ism vs. certain sophisticated versions of presentism.)

DDG

The question "Is a beginning or end in time possible?" gives one an unshakeable sense of metaphysical mystery.- of something far beyond our understanding. And there is truth in that - there *is* a mystery, but the world does not *give it to us* in these terms. 'The world has an end.' 'The world is endless.' We are just animals making noises. These noises - if they mean anything - must mean something highly singular - i.e. we have our concepts of a process ending, starting, running indefinitely... all this presupposes time. Then we shift to talking of time itself as a process - as though time, while not a process, were enough like a process for us to speak this way via analogy. But that is nonsense.

'The tense-logician must not pre-judge the structure of time, since this is a job for physics. Hence the alternative systems - sets of axioms and corresponding classes of frames.'. But has physics ever settled anything like that before? That is: on what ground does one call it a job for physics? What about 'a job for physics or metaphysics'? Well, that takes the heat off the statement - or rather, dissipates it - but now our talk of 'settling things' has become ridiculous.

DDG

* * *

I don't want to give an *analysis* of modality in terms of concepts. Obviously, 'bachelors are unmarried' is not the same as any proposition *about* the concept of bachelorhood, and to overlook this will keep one firmly in the dark.

"Truth-making relations" come in all shapes and sizes.

Secondly, necessities may be grounded by concepts (in a sense), but concepts are often in turn grounded (in another sense) by reality, by experience. And so it's not all about "truth-making" here.

"Adequacy-making". Adequacy-makers would be the sorts of things Wittgenstein was, in his middle period, urging that we can't talk about. But my take on it, and perhaps Wittgenstein's later take, is that we often *can* talk about such things, although we are tempted to read into this a quasi-causal, or an explanatory, relation. I.e., we *can* say that our conceptual situation with respect to Venus - the master concept, and the two merged aspect-concepts - is made adequate by the fact that there is this one object Venus, which we view in two ways, etc. We have just one concept here, because there is just one object - but anyone who is not blind will sense that there is something, so to speak, quasi-tautologous in this. It seems to amount to 'We have one just one concept here, and rightly' (but of course we mustn't say it literally *means* that).

It might therefore seem that all my talk of adequacy of concepts etc. with respect to modality is empty. I have no objection to its being called empty. I am trying to show (and make) conceptual connections.

* * *

What direction your concepts should develop in is not a purely internal matter. And that is the key to why the determination of necessity cannot be a purely internal matter.

* * *

If a thought is internally impossible, we can work this out if we're smart enough.

If a thought is merely externally impossible, we cannot.

There is a crudity lurking in this, however. I believe it consists in this: suppose we presently believe that the hooded killer and the butler are different people, but know enough relevant facts about the whole case - the role of the other people involved, etc. - and are in a position to work out that they are the same person, Sherlock Holmes style. In many cases like that we would want to say that these relevant facts clearly do not enter into the essence (identity conditions) of the thought that the hooded killer is not the butler. And so the false thought (individuated narrowly and internally) that the hooded killer is distinct from the butler, although internally impossible according to the first naive conditional above, could hold good in different cognitive (internal) surroundings, let alone different externals.

((This is connected with the Twin-Earth-like cases in the realm of the a priori.))

* * *

Is this what I want to say?: Once you've got the truth value, the modal status is always a priori. For the necessary aposteriori, at any rate, this seems to hold.

"In all cases where one needs empirical input to decide on modal status, one also needs empirical input to decide on truth-value."

* * *

When talking about external modality, we are in an important sense *not* hostage to the possibility of a vastly "better" scheme which does not recognize objects we take to exist. We may put it this way: all we can do is play our game, and try to improve and extend it. (This relates to some of David Lewis's methodological points, except he's more about simplification and unification. Neurath's Boat, too.) We might have doubts about some of our games, and perhaps be attracted to instrumentalism. Fine. But we still think. And in a sense that commits us to the acceptability of a conceptual scheme involving such things as we think of. In moments where we doubt things, however, then we doubt them, and that's all there is to it. (Postmodernism.)

* * *

"You can't have an inferior *version* of a particular concept - that makes no sense. If you have the concept, you have it. If you don't enjoy this distinction, well, then you either have another concept in its place, or nothing which corresponds at all."

We can see this as a kind of grammatical prescription. And it's very hard to see immediately *why* it would not be a good one to follow. It may seem attractively simple, clear-cut, systematic. But remember: we do not have a thorough overview, ready-to-hand, of our use of 'concept', and so for all we know right off the bat, a thoroughgoing adoption of this prescription may be quite crippling.

In such situations, we have to stay calm above all, and not do anything rash. (There is a place for what might be called philosophical faith, but not for philosophical *leaps of faith*.)

* * *

Counterfactuals and might-have-beens as non-factive. At first this seems like something very solid: there is the way things are - that constitutes the facts - and then there are the ways things could have been, would have been if ..., etc. This gives us the picture of a solid body surrounded by a kind of atmosphere, perhaps vaporous or shadowy, produced by the body. But of course we can also say that things *are* such that <some

counterfactual>, <some metaphysical modal statement>, etc.

* * *

I need to talk about 'modal character' of beliefs without endorsing them - I need to be able to call beliefs, thoughts, sentences something like contingent, and something like necessary, but without implying that they are actually true. To this end, 'contingent character' and 'necessary character'.

* * *

It does not seem impossible in principle for us to meet beings so smart that those of us who were able to learn their ways would come to the view that they themselves used to understand and know very little - were perhaps under all kinds of illusions, etc.

But then, if I really think this is quite possible, shouldn't I draw the moral now and become a good deal more circumspect about the way I think?

But there is nowhere for me to go with it. And of course that doesn't mean I should conclude from this that I am justified in thinking as I do.

'Coming to believe X would leave me high and dry, hence I am right to reject it.'

What is being appealed to here? Practicality? Morality? Some cognitive norm?

* * *

The difference between necessary and contingent appears to us almost as a difference in the nature of the signs. And when a "method of projection" is included in the identity of the sign, then this is indeed the case.

* * *

Is this the way forward?:

Forget trying to tell any story about what makes metaphysical necessities *true*. This is just one of the problematic aspects of truth.

Instead, concentrate on the difference between necessary and contingent *character*.

The rest of what I've been saying could be seen as a restricted account of truth:

If p is of necessary character, then it is true iff ...

In the gap might go something about adequacy of concepts involved. Of course, this might not be very illuminating.

This brings out a problem: how can modal character come in advance? Don't we need the right concepts for that?

What I mean is: the stuff above may give the impression that I regard modal character as something which is transparent, or at any rate which can be settled prior to settling the truth-value. But this appears not to be the case in general.

This example was the first which came to mind:

'John is in the same place as Harry' - when John = Harry, necessary, otherwise contingent.

Now, this example seems interestingly artificial; we wouldn't usually say this if we knew that John = Harry, except perhaps as a kind of Socratic joke - for example, in a conversation with two other people, where one knows the identity and the other doesn't, you and the other knower may have a kind of private joke by telling the other person this).

Are there less artificial examples - i.e. ones yielding on both sides utterances which might represent more practical, sincere thoughts? (Not that I am hoping for this or anything.)

If John goes to the party, Harry goes too.

Another kind of potential case, seemingly less artificial:

'Water has no characteristic internal structure'.

With an inadequate concept of water, this might look contingently true. With an adequate concept, necessarily false.

This is a very interesting case, actually. Like the puzzle of the contingent apriori, it shows how much turns on identity conditions for propositions, thoughts and concepts. For, in one sense, the very same thought which was of contingent character, after empirically due conceptual change, might become necessary in character. However, in another sense, these aren't the same thought. Indeed, in the new system, a quite different construction is actually modally closer to the original thought in the old system than the new necessity - namely, 'Watery stuff has no characteristic internal structure', or some longer proposition involving the old 'clear, potable, falls from the sky, found in oceans and lakes' description which 2D semanticists know off by heart.

* * *

At the beginning of Nietzsche's Zarathustra, the madman proclaiming the death of God is jeered at by the people, who don't believe in Him and do not consider this to be sensible news. Nietzsche believed that the important philosophical consequences of "the death of God" had yet to be drawn. This idea illuminates the history of the philosophy of modality in the 20th century - especially if one compares with what Leibniz and Spinoza had to say on the matter. (And, less dramatically, Kant, Descartes and less notably, Wolff.)

Put simply, the history of analytic philosophy of modality can be seen as the story of the drawing of the consequence that, without God - in short, without the notion of a perfect mind - we've got a serious gap in our old accounting system, and must rethink the whole thing. (And needless to say, simply putting the notion of a perfect mind back in is no way forward.)

It was thought that Humankind was made in God's image. We learnt more about the former, and this became untenable.

Or: God was reckoned in Humankind's image, but when we learnt more about the latter, the projection became untenable.

(With respect to the demise of theism, logical positivism stands to modality as utilitarianism stands to morality. Lifeless, abstract, over-simple.)

The primary task is to come to a philosophical understanding of modality without using the notion of the mind of God. Secondly, it can and should be made apparent that even if we do countenance this notion, we wind up in full-blown theology with all its problems and sterility in the face of modern knowledge. We can't, for instance, use a 'cleaned up' notion of the mind of God as a "purely theoretical posit". Any such thing would raise a thousand unanswerable questions.

In terms of 'having a philosophical account of modality', Leibniz and Spinoza might be regarded as the high points in the history of thought. They had more to say in answer to the most fundamental problems than anyone else I know of.

In some sense, I have to go one better. But this does not mean that I have to give "more of an account". In many ways I want *less* of an account - I don't want to (re-)ground modal discourse. Rather, 'going one better' here means: a more thorough, better-informed consideration of the matter. Or a better approach. (This is slightly daunting, given that one of these guys invented the differential calculus. But that's an anachronistic worry.)

Of course, Leibniz and Spinoza didn't have the problems of the 20th century in view - the Kripkean problems, externalism, etc. - but the point is that they would have been able to handle them in a way we

can't.

* * *

The question of the mind of God has become unavoidable for me in this work. This was most unanticipated, and it raises difficult authorial problems. For I do not want to write a book which denies (or asserts) the existence of God. And this raises the further problem of what to say about what I believe - for I cannot very well remain silent on that.

Do I believe in God? The question is not an easy one. At any rate, it seems clear to be that I don't have any kind of positive faith which is relevant here. So I need to make this clear, but without coming off as one of those narrow, smug atheists. (Perhaps this reflection itself makes things clear enough.)

The Necessary Aposteriori

The pre-astronomical concepts of Hesperus as distinct from Phosphorus - internally fine in themselves, but given that both are *of* Venus, while being regarded as being of distinct objects, the propositions of necessary character which are issued in by them are metaphysically necessarily false.

There are also cases where the stuff I was calling external (facts about what these empirical object-concepts are *of*) is "absorbed" into the scheme, lying dormant internally - widely scattered, so that one may be in a position to 'put two and two together'.

Imagine Lois Lane was in that kind of situation - i.e., knew enough facts to work out that Clark Kent is Superman, if only she reflected in the right manner. Now, her conceptual scheme as a whole seems to internally point toward amalgamating the Clark Kent and Superman concepts. But it seems like this "dissonance" is not harboured within her thought that Clark Kent and Superman are distinct - it is not that localized.

All Descartes' criteria for modal evaluation are "internal" - clearness, distinctness, freedom from contradiction. But - and this is most clear with respect to individual concepts and the Kripkean necessary aposteriori - some of our concepts are only faulty in that they are inadequate to their objects. I.e. external facts make it the case that they have the objects they have, and given that, they may fail. But they would be adequate to other possible - even actual (Twin Earth) - objects..

I will have to discuss how this relates to other sorts of concepts.

It is essential to make enough room for "coherent misconceptions". I have first become clear about this with respect to individuals - the clearest cases of the necessary a posteriori. Perhaps also water and the like. But there also seems to be a place for such considerations with respect to other sorts of concepts. This is important, because it shows how modality is not arbitrary, not 'merely conventional' in any sense.

Concepts are not like words - there is nothing arbitrary or conventional about them in any strong sense. (Although of course there can be two different ways of thinking about some matter, where it is in some sense arbitrary which one is adopted.)

The external-conceptual viewpoint on modality does not come to any abrupt halt. It is extremely strong for 'Hesperus is Phosphorus' and the like, but also has application elsewhere, just like other 'objectual' forms. The concept of happiness has happiness as its object. The concept of freedom, freedom. (The language-game can be employed in this region too, but that doesn't mean it's not a very different region.)

"Modal intuitions" as 'concept-world' judgements: judgements about what the world calls for with respect to concepts - modulo sophistication over 'about' here. For all its crudeness, this is preferable at least to the inappropriate picture of a *hunch* about some matter of fact, as it were an inner voice which says 'The modal

situation is this ...'. One at least gets some kind of traction - analogies with designing a machine, a computer program, etc.

* * *

In what sense are 'Hesperus is Phosphorus' and 'Water is H₂O' true in virtue of meaning, or concepts?

Here 'truth in virtue of' starts to seem like the wrong formulation. One wants to say: they are not just true in virtue of concepts - the story doesn't end there.

The Kripkean examples of the necessary a posteriori straddle form and content.

* * *

I see a moving dot in the distance, and as it gets closer, I don't know whether it's a car or a horse. I have minted an individual concept for this thing I'm looking at, which is developing as the thing approaches. There may even be an initial stage where I don't know whether the 'object' is a physical body or a visual disturbance, but suppose I soon go for the first hypothesis - then the concept develops. The object is far away though, and could either be a car or a horse... but once it approaches, I see it is a horse. My concept of this individual develops further - and now I am not prepared to say ("counterfactually de re") that that very horse could have been a car. Thus now the thought 'this very thing is a car' is internally impossible. But before, when I knew less, I could have internally coherently thought 'this very thing is a car'. So that latter thought was internally possible, but given what the individual concept involved was *of*, we now recognize that the concept involved was inadequate to its object - and given that an adequate conception of the object internally rules out its being a car, the thought was not *externally* (metaphysically) possible.

I initially formulated the moral as follows: external possibility is internal possibility plus adequacy of the involved concepts to their objects.

But that's not right... for you don't actually have to have adequate concepts to hit on an external possibility. (The import of this is something I will no doubt be occupied with later.)

Instead: an externally possible (necessary) thought is one which, when refitted (if need be) with externally adequate concepts, is internally possible (necessary). (This is meant to be trivial in a sense. Insofar as it is a definition, it is a definition of 'externally adequate' rather than 'externally possible (necessary)').

Thus if Lois Lane, with her inadequate (because separate) concepts of Sup and CK, thinks they are distinct, her thought is internally possible, and even internally necessary. In *her* system, the thought that they are identical is internally impossible, even though it's externally necessary: substitute in proper concepts (the Sup and CK concepts would in that case appear as aspect-concepts belonging to a common master) and you get an internal necessity.

I think that goes some way to making sense of 'the necessary a posteriori'.

* * *

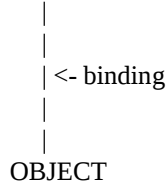
Given the external content of a thought - what "out there" it is of - the concepts involved may be in themselves perfectly in order, i.e. may be adequate for some possible objects, while nevertheless inadequate to their actual objects.

An indistinct worry: when I say that some particular concept is inadequate to its object - I can only have that thought by employing my *own* concept of that object. (But that doesn't mean I presuppose that my own concept *is* adequate - I might question that, but I also have to use it to select the object to which it is inadequate. (Dual supposition.))

When talking about some concept which is someone else's concept of an object we know, or not our main concept, an easy mistake would be to employ the picture of comparing two concepts side by side. But that's not what we're doing - for we are talking *about* the inadequate concept, and for that we employ our concept

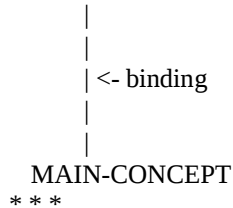
of that concept. My concept of my concept of an object is *bound* to my concept of that object - whereas I might also have a concept of another (foreign or subsidiary) concept of the same object, which does not - so to speak - cover the object for me (e.g. the concept of Hesperus as the evening star). Thus while I can compare these concepts "side by side", there is another *dimension*:

MAIN-CONCEPT----coextension----INADEQUATE-CONCEPT



Or to label the situation from an external standpoint (i.e. where every 'node' is a concept):

META-MAIN-CONCEPT----metacoextension----META-INADEQUATE-CONCEPT



Kripke's examples of the necessary a posteriori drive a wedge between, not just modality and "the a priori", but also modality and conventionality, arbitrariness. (The fall of logical positivism.) And one must keep pushing.

For once we see that our concepts of individuals and kinds may not be adequate to their objects, we see that this can happen quite widely: one may, for example, not have an adequate concept of freedom, or money, even if one's present concept does not harbour any sort of 'internal inconsistency' - i.e. if one's present concept is an inadequate tool, rather than an out-of-order version of the adequate one. The inadequacy in such a case, I want to say, is external: it has to do with how life is, what the world is like, and the role of the concepts involved.

When we have an inadequate but coherent concept of something, what makes it an inadequate concept of *that thing*, rather than an adequate concept of something else? One cannot simply give some answer which appears to work for concepts of empirically observable individuals, substances and the like. What about cases such as: freedom, arthritis, love, happiness, pulleys, long division, parliament, computation, consciousness, relative motion, time, sarcasm, the corpuscularian philosophy, and jokes? These all seem to be things which one can in some sense have an inadequate (or primitive) concept of.

This whole area is confusing because we oscillate between thinking of concepts as internally individuated - by their structure and position in the wider structure, so to speak - and externally.

* * *

'Names of empirical things are indexical' - if a philosopher said this in the context of a discussion of modality and semantic internalism/externalism, I would understand perfectly.

Indexical in this context = Twin Earthable.

Suppose, unbeknownst to us, Earth has a number of "twins" which run in perfect harmony with ours. Thus someone who knew this, and who was blindfolded and placed on one of them at random, would be in a position where it would be most natural for them to say: 'I don't know who these people are, which ones they are.'

And yet someone who knew nothing about the twin Earths, we would usually say, knows perfectly well who they're talking to, etc.

Could we put this as so?: sometimes an increase in qualitative knowledge leads to a decrease in haeccectic knowledge. But this does not seem to be a proper answer. Such paradoxical formulations always miss the moral of the story.

If such a person were planted on one of these identical planets and forced to live there for an extended period, it would become natural again for them to talk knowingly about 'who people are'.

The confusing thing is to understand the sense in which 'knowing who someone is' is a relative matter.

Take the unsuspecting Earthian: we want to say that, despite it being true that if he woke up on a twin Earth tomorrow, he would misidentify everyone as those he knows, he *is* able to identify the ones he knows.

It seems like we could say there are two uses of 'Who is this?', 'Who are you?', etc. Or rather, two ways of answering it. By merging two aspects, or by providing a new aspect.

* * *

We make this comparison: 'Things could have been thus' -> 'Our system can be configured thus, and is adequate in this respect'.

At first glance, it might look like this way of looking at things is refuted by cases like this:

'Things could have been such as to be unrepresentable in our system' or to be clearer about scope: 'Our system is such that things could have been such as to be unrepresentable in it'.

In fact, there is no conflict here; if those things are true, one can say: this is because our system can be configured to say 'There are situations not representable in this system', and is adequate in this respect.

(A similar point holds for the apparent problem of the possibility of inconceivable objects.)

But now comes the question whether it is, or could be, adequate for a system (or our system) to allow this.

Now it might look again like we are up against strange, ultra-general problems whose solutions are inaccessible.

'Is an ur-system of representation possible?' (Ur-system - umbrella system, cover-all system.)

Perhaps all we need is the ability to talk of other systems and configurations, and then anything they can say can be said by us thus: 'Things are as represented by this configuration in that system'.

(Reflecting on this makes it seem plausible that, if we can't fully describe something, then neither can we fully describe a descriptive system which can - for if we could, we could then say that things are how they are when that descriptive system is rightly in THIS configuration.)

It can easily look as though, if we are going to talk of indescribable things which might have happened, then these things must really (for our talk to be right) be indescribable, in any system of representation. Otherwise, it is merely a paucity in our system as it is for us. And so it can come to look as though an ideal-for-actuality system of representation must be capable of describing any metaphysically possible scenario.

(Paucity, but justifiable? Stuff to follow up here, beyond me right now.)

It can seem like any kind of "maximally adequate" scheme *must* have all the power of less adequate

schemes, since one must after all be able to describe what is done with them. I.e., describing necessarily false beliefs, describing what is said in a less adequate language. Or: perhaps this just can't fully be done. (Perhaps omniscience is in this way impossible. God might not be able to *give the content* of inadequate thoughts, in the sense that one needs inadequate concepts to do that.)

This is of course not the same as saying that perhaps a "maximally adequate" scheme will fall short of representing all possibilities. It may be that way too (if there are multiple good schemes which can't be reconciled), but here the issue is whether it may also be that the representations unavailable in the maximal scheme only exist in other schemes due to their inadequacy. If this were the case, then perhaps in order to accurately *give the content* of certain propositions and intentional states, one would need concepts which one really ought not to have, for other reasons. So, one shouldn't have such concepts: and yet, if one does, one can do stuff which can't be done otherwise.

Is this paradoxical, or just unexpected? (This itself seems to be an instance of the paradox! Setting up the paradox might require the setting up of an inadequate conceptual situation. And this may in a sense be requisite for 'an understanding of the paradox'. In this connection, it has often struck me as quite plausible that it is easier to understand and recall certain philosophical ideas and works if one shares certain confusions or false views with their creators.)

The notion of simulation (of conceptual situations and thoughts) is relevant here.

Another puzzle: I have said that epistemic modality is open, subjunctive closed. Hesperus could be Phosphorus, but given that it's not, it couldn't have been distinct from Phosphorus. Thus it looks like when evaluating metaphysical (subjunctive) modals, we fix on our endorsed scheme. (Here schemes are being individuated in a very fine-grained manner; change your views about an identity involving proper names, and you thereby move to another scheme.) However - and this is the puzzle - it seems that we can say: We could have thought Hesperus and Phosphorus were distinct. We could have thought that water was XYZ.

Here it might look like we are relaxing the fixedness of our endorsed scheme, in order to represent the contents of these thoughts - since we can't represent those contents directly in our specific, endorsed scheme.

It is tempting (and may be right) to say: no, things are just as closed here - these cases are metaconceptual.

Two putative states of affairs: 'S thinks Hesperus isn't Phosphorus', 'Hesperus isn't Phosphorus'. The first is metaphysically possible, the second is not. Thus I want to say: the first can be directly represented in our endorsed system, the second cannot.

Should we then say that, in thinking of S's thinking that Hesperus isn't Phosphorus, we are simulating S's thought? Simulating another system? That seems gratuitous. Some kind of 'metaconceptual' or 'metasemantic' move, some kind of going-up-a-level, seems more apt here.

In our system, taken as fixed, we do not have a representation corresponding to 'Hesperus is not Phosphorus', as do possible inadequate systems. However, we *do* have the concept of the proposition that Hesperus is Phosphorus, and we have the concepts of the concepts of Hesperus and Phosphorus, so we can conceive of the "foreign" proposition that Hesperus is not Phosphorus. Or: we can conceive of this representation, in our endorsed system - but it is not a representation in our system. This is all that is needed.

But: these necessary *a posteriori* claims are perhaps quite a special case. How, then, does the account just indicated fare with respect to representations more remote from anything in our system?

Above, it was said that we use the concept of the concept of Hesperus, and the concept of the concept of Phosphorus, and so conceive of the metaphysically impossible representation. However, this case also has the feature that we have the concepts of Hesperus and Phosphorus themselves (or 'mutually unified' versions, in the form of aspect-concepts). In other cases, we might be working with the concept of a

concept which we have no version of.

There is a lot of difference between the two kinds of cases.

Three levels of conceptual involvement: representable in system, metarepresentable using concepts of [concepts which we have versions of in our system], and finally metarepresentable using concepts of [concepts which have no counterparts in our system].

(The square brackets above disambiguate with respect to scope.)

In all that I've written so far, I haven't dwelt much on the possibility of quite different conceptual systems - as one is led to, for example, by Wittgenstein in Remarks on the Phil. of Psych. Vol. 2, s197-8. A world where - as we'd put it - the colours of things were constantly in flux, no one ever saw a fixed colour.)

I can imagine someone reading what I've written so far and thinking that I am being dogmatic in this connection (in some vague way).

'A maximally adequate conceptual scheme' - in most cases where one wants to talk about such a thing, it is usually something we could also call 'a maximally adequate version of our scheme'. Or perhaps: 'that maximally adequate scheme which lies closest to our present one'.

One can describe certain environments, containing thinking beings, and can often then catch a glimpse into a quite different conceptual scheme. (I.e. imagining different certain empirical facts to be otherwise, facts which are so familiar to us that we do not think of them. E.g. that persons always have just one body.)

But then, every time we do this, couldn't we imagine a kind of highly dynamic agent which somehow wears both hats - i.e. perhaps a being which can undergo bodily transformation, and make sense of an environment like ours, as well as these others? If so, then isn't the maximally adequate scheme the one which as it were subsumes all possible adequate concepts whatever?

What I'm staring at here could be called 'the ur-scheme problem'.

In some cases where we wonder at the (metaphysical, conceptual) possibility of something, imaginability, conceivability *in full detail* seems to be close to what is at issue - i.e., whether the thing can be 'worked out'. In other cases however, it is quite different. In other cases, the question is closer to: *ought* we conceive of this? (To see how the former shades into the latter would be illuminating. And this might have to do with the nature of the task of 'working out' in some particular case, and the extent to which this task is "clear in advance". ((This is connected with (what I want to call) railed vs. unrailed spontaneous concept formation in Wittgenstein's philosophy of math, knowing how to go on vs. discovering how.)))

An example of the latter: suppose someone says that when a cat weaves around someone's legs, they are asserting ownership of this person. (I was actually told this as a child.) We might now consider whether this is so, get puzzled, and then consider whether it is even possible for part of a normal cat's normal behaviour to turn out to amount to an assertion of ownership.) Here, it is quite unnatural to think that any kind of conceivability or imaginability (however detailed) is at issue. We have seen the behaviour a hundred times. We can imagine *it* perfectly well. What, now, does it mean to imagine it being an assertion of ownership? (Furthermore, *actually* being such an assertion - by which I mean, it does not count to merely imagine that one is dealing with some extraordinary cat who can think and communicate with conventional signals, or indeed anything out of the ordinary.)

The assertion that the cat is asserting ownership, and countless others like it, have a special character - it is granted in advance, so to speak, that the cat is not asserting ownership by the *usual* criteria. Conceptually progressive propositions, but also linguistically progressive propositions. (The former are perhaps well seen as a proper subset of the latter.)

One feels like saying: here, a decision is needed. (Although this is misleading in that no *conscious* decision *is* needed, no freedom to decide either way need be felt, etc.)

Within such cases, one might distinguish those which as it were call for full-blown (literal?) acceptance, while others are not concerned with that, being, rather, quasi-metaphorical, quasi-analogies. But - and hopefully this seems trite - there is no clear line between literal truth and analogy, figure, simile, metaphor. However, with the latter, one of the characteristic things is that a decision isn't really needed in the full sense. However, one may still reject or embrace ('work with') the idea, criticize it, modify it, etc. ((Simile as a minimal literalization of metaphor - in effect, the literal truth one gets to when one accepts a metaphor as a metaphor. Which isn't to say that all similes are such literalizations, but rather: given an acceptable metaphor, one can get a literally true simile - at least, that's the idea.))

An interesting feature of the cat case is that, in a sense, the verification and falsification conditions are already in place. It is already clear that, if one accepts this kind of proposition at all, one will count 'the cat asserted ownership' true when the cat weaved around the person's legs, false otherwise. Obviously it will be a more subtle than that, for many reasons - vagueness about whether weaving took place, the possibility of the cat being in some pathological state which just happens to involve weaving, etc. - but still, the point remains that the *working* criteria are not the focus of the problem. Reflection on this leads to the comparison of the cat proposition (in its progressive use) with a proposal of the adoption of a norm of expression.

Whenever something like that is said (i.e. 'a norm of expression'), it looks as though something important is being skirted over too easily. This is connected with the fact that this is no merely conventional, arbitrary norm of expression. It is highly charged with significance, with meaning. Also: to say it is a norm of expression might suggest that it is *not* a norm of thought. That thought goes on underneath, and this norm relates to how the thought gets 'put into words'. But the cat case is obviously not like that.

To think of a badly treated mechanical device as suffering, feeling sympathy for it - this is a perfectly possible form of thought which, as a matter of fact, ordinary adults do not engage in. (Here my saying 'as a matter of fact' may give a wrong impression, namely that I think it would be 'just as good' to do that, or that we can't say anything against such a viewpoint, etc. But not at all. I'm not getting into that.)

The cat case, and the case of the suffering vacuum cleaner, are interesting partly because they make trouble for a certain oversimplification of the working of language which is, in certain circumstances, natural for us.

The oversimplification is something like this: in cases where the working criteria are not an issue, but where judgement one way or the other sets a kind of precedent (for the remainder of a conversation or train of thought (internal monologue), at least), it may look like the issue is 'merely verbal', 'merely pragmatic', or something of the sort. But this is liable to be extremely misleading; it focuses on the comparison of such a case to that of, say, the adoption of a perfectly arbitrary label, to the exclusion of other comparisons. The *consequences* are very different here from those which follow an arbitrary labelling decision. (Associations.)

It is instructive to reflect that sometimes we accept provisionally a conceptual precedent for the purposes of a conversation (or even a private train of thought), but have misgivings about it. We feel, as it were, that quite possibly we can get on OK with it for now, that to start critiquing here might be impractical at this point (e.g. in an involved conversation where we will have to go to bed soon, or get off the phone or something), but that, were the stakes to be raised, so to speak, we would have to try to sort it out.

Consider a conversation of this sort. The limitations one characteristically feels in such discussions. (Consider especially the case where your conversation partner doesn't see any difficulty.) Thoughts have potential problems other than falsity! Truths, expressed in a certain way, can be almost false. One almost wants to talk here of a 'higher' kind of truth and falsity, of verification and falsification. (I think Hegel actually does something of this sort at the beginning of his Logic.)

'They believe a machine suffers!' - this could be thought to be an inadequate, misleading statement in a similar way to 'Cantor showed that there are numbers larger than the number of natural numbers' - except in the first case, the thing is rejected as absurd, in the second case, accepted as mind-bogglingly wonderful. (There are, of course, reasons for these different treatments.) In both cases, an extended conception is being regarded from the point of view of a narrower one.

((There is a false note here, interestingly. I feel like saying: no, those who believe that the machine suffers do not have an extended concept, it is we who have a restricted one. We've got more structure, so to speak. We could have a concept similar to but distinct from our concept of suffering, which we apply to machines. Thus we have a sort of distinction where they have none. Or, we could apply something like that more minimal concept across the board, and then add something for the case of "real suffering".))

The proposition that a machine suffers could actually be embedded, quite deeply, into a highly sophisticated conceptual framework. The point being: it is not a proposition which can only survive in very primitive or childish systems. One could talk of panprotopsychism (Chalmers), a continuum of organizational complexity, and with it a continuum of suffering. So complex human computers in disorder, on such a view, perhaps suffer about as much as an insect. How far can it be taken? To vacuum cleaners? Hot water bottles? Not the latter. (The human teleology (purpose) of the object, if such there be, would naturally not play any part in its placement on the continuum.)

How do we know that rocks aren't incredibly sad? If this is metaphysically impossible, how do we know *that*?

(This is another case where imagination seems quite irrelevant. There's nothing to picture except the rock. Or, one can imagine being sad. Neither procedure constitutes an engagement with the case, however.)

I believe that any reason given to think that rocks aren't sad is going to be bullshit. What is disturbing here is the strong inclination toward finding such reasons nonetheless. It suggests deep infelicities in our thinking.

The Contingent Apriori

(The following paragraph is toward a solution of the problems associated with the contingent a priori.) With both time and spatial length measurements, as well as colour-naming, there is a sense in which our standard is not fully contained in any particular sample. This makes the grammar of statements of measurement complex and hard to survey, especially when we talk about the length of standard samples. Despite the fact that we actually do not have one single court of appeal for these statements, we go on trying to use the picture of one paradigm, one single object of comparison, and thus arises the notion of the length itself as an ethereal object. (I am not for a moment suggesting that we should stop using this picture, nor that we should continue - I am not saying anything of that sort, but merely trying to describe.)

* * *

'This stick could have been another length.' - the stick itself comes into this proposition twice, as it were: in 'this stick', and to give sense to 'another length'. For 'another length' only makes sense in relation to 'this length'.

* * *

Samples have a singular role. Part of the representational apparatus, but not like a word in that they only 'work' because of specific qualities they have, and of course not like a concept. When we rely on a sample in order to use propositions about some particular colour, it can look as though what is happening is that our concept of that colour is not adequate or complete in itself, but needs the sample to prop it up - as a kind of prosthetic aid.

* * *

EDIT

"How could it be both contingent and apriori that the stick is 1m long?! Suppose it is contingent: that

means it could have been less than 1m long. Now, if you were in that world - how the hell would you be able to work out apriori that it is a metre long?"

About a million things have gone wrong here.

For a start, it's not a matter of not being able to work out, in that world, that the metre stick is 1m long. For, in that world, it isn't!

Talk instead about the proposition that the metre stick is as long as the metre stick, and then you get two readings - the non-rigid and the rigid (with respect to the length).

Evans, Julius and the zip.

Descriptive names.

Burrowing down, detachability. Two readings of 'the metre stick is as long as the metre stick': shallow and deep. With shallow, no independent, detachable length-concept is reached. With deep, it is.

The Russell yacht case. With the initial remark - "I thought your yacht would be bigger than it is" - the 'it is' burrows down to the visually-formed, well-connected concept of the yacht's actual length. With the reply - 'No, my yacht is not bigger than it is' - the 'it is' just invokes the shallow, unconnected concept - as it were the definite description, the flaccid designator.

In Kripkean terms, the problem comes from the fact measurement expressions are like rigid designators of lengths.

And we fix the reference with a non-essential feature.

"I know that someone invented the zip. I hereby stipulate that 'Julius' refers to that person. Now, concerning Julius, I know that this very person invented the zip.' So now this subject can have a thought of the form 'that very person invented the zip'.

But suppose they are at party one day and someone brings someone to the subject, points at them, and says 'this person invented the zip'. The subject did not know this. Thus there is not just *one* possible thought.

If someone hasn't seen the metre stick, or hasn't seen the inventor of the zip, there is a sense that when they make their stipulation, they are getting an individual concept, a rigid concept, but as it were on lay by. They've got the thing ready, because they know something answers to it, but as of yet they can't do anything independent with it.

"There must be a solution to this problem; let us call it S. Now we know the solution: it is S." - One can talk that way, to be sure. But where does it get you?

(We can bring the Evans-like namey cases a little closer to the metre-case or the color-sample case by talking about, say, someone person's hat. 'The metre stick might have had a different length', 'This guy might have had a different hat'.)

A rigid concept is formed. It seems like the central thing, w.r.t. to the problem of the contingent apriori, to see that the very identity conditions of this concept include both its definitional connection, and the identity of its object.

Make that grammatical stipulation, at any rate. Now, it is clear that it is logically impossible for someone in one of the falsifying worlds to have that concept.

That is pretty inarticulate still, but I think it constitutes the solution.

(The problem of the contingent apriori, I am now hoping, was solved (about 3 mins ago) at approx. 5:17pm, 13th Feb 2011.)

Individuate by definition alone, and it's not contingent. Individuate by object alone, and it's not bound to be apriori.

OK, so what does this solution mean for the Russian doll picture of modality?

Well, firstly, the solution forces us to remember that there is a way of structuring thoughts such that they are essentially not true of actuality - such that their non-actuality is a priori.

That is: some thoughts are counterfactual by their very nature, their internal nature; a counterfactual thought is not simply a false thought. Thinking of the world as being different *per se* from the way it actually is.

Suppose Tower A is 15m tall:

Accidentally counterfactual thought: the thought of Tower A being 30m tall.

Essentially counterfactual thought: the thought of Tower A being twice as tall as it actually is.

The same, contingently non-actual, state of affairs answers to both. But the falsity of the first is aposteriori, the falsity of the second apriori.

The proposition that Tower A is twice as tall as it actually is, is plainly contradictory, its negation a plain tautology.

But if you now introduce a sign X for the a unit of length the standard of which is Tower A, you get the pair:

'Tower A is 2X tall', 'Tower A is not 2X tall'.

Now, it can be asked, do the thoughts represented by *these* propositions possess essential counterfactuality and actuality respectively, or accidental? And the answer is that that obviously depends on how you individuate thoughts. There is no single natural way, here. If identity of logical form and identity of reference (broadly construed to include measurement expressions) is taken as sufficient, then it is not part of the essence of the thought that it is apriori - only one special version of it would be, so to speak. (We may think here of equivalence classes of possible thoughts-individuated-in-a-finer-grained-way.) But if identity of sense is part of the identity conditions for thoughts, then the propositions possess essential counterfactuality and actuality. Now, here I include in 'sense' definitional connections. Kripke makes a big deal about the distinction between giving the sense and fixing the reference. But he doesn't in his own stuff talk about the senses of names anyway (I should check this). Now, the distinction is of course important with respect to rigid designation. 'Julius' doesn't really *mean* 'the inventor of the zip', although it is - one could say - part of the meaning of 'Julius' that its referent actually invented the zip. Kripke distinguishes between two kinds of definitions here. This is a good distinction, but one must realize that even in the reference-fixing case of definition, the sense of the defined term can be said to involve this definitional connection. It is part of the sense of the name that its *referent* (actual referent) is the same as the reference-fixing definition. Sense of the name, individual concept tied to the name - one can use various expressions here. (That is to say: I'm not trying to be dogmatic or normative here with respect to the word 'sense' - I just felt like using it that way.)

The

For a thought in the relevant sense to be contingent in the relevant sense does not mean: that very thought could be in an alternative situation where it is a false thought, but rather: the actual system permits one to negate the thought, qua counterfactual - within the system as it is (and so not like the negation of the thought that Hesperus is Phosphorus and other empirical necessities: they have intelligible negations, but

these have their place in a modified conceptual system).

14.2.11.

If there is a stick or a colour sample which one hasn't seen, one can still stipulate that the unit length expression '1u' is to refer to the actual, present length of the stick, and that the expression 'C' is to be a name of the actual, present colour of the sample. Now, one can say that the stick is 1u long, and that the sample is C in colour.

Plainly, these statements do not embody any knowledge beyond the knowledge that there is such and such a stick and such and such a sample, and (in another sense) that these definitions are in place. But if we got acquainted with the stick and the colour sample, started using the 'u' measuring system and the colour name 'C' to describe other objects, the game has changed. We get an idea of the length 'u' and the colour C which we can carry around with us. And of course there will be circumstances in which we would say that the stick is no longer 1u long, and the sample is no longer C. (In the verification of u- and C-statements, there is no single highest court of appeal, in Wittgenstein's phrase.)

Consider the fact that the definitions of 'u' and 'C' are not saying that, in the future, the length of the stick and the colour of the sample are to be called '1u' and 'C' - rather, reference is being made to a specific length, via an object as it is at a time. Thus, when using the expressions defined, the criteria for statements involving them cannot not simply involve correlation with a sample. No, if correlation with these samples is to be spoken of at all here, then it must also be verified that the length and colour of the samples has not changed. And how is this to be done? Obviously, not by means of the samples.

It is not apriori at all that the metre stick is now 1m long. Only, that it was 1m long at the time referred to in the definition of 'm'.

There is of course another language game we could play, where correlation with the sample really is the verification, and we

If a sample is relied upon for the expression E, and then goes out of existence, E-statements acquire something of the character of counterfactual conditionals.

* * *

The appearance of contingent a priori knowledge is generated only if we fail to distinguish these two salient truths, mistakenly running together the *apriority* of the first and the *contingency* of the second. You could of course know that *s* is one meter long at time *t*, by virtue of knowing (a) that any stick used to fix the reference of a unit of measurement will measure exactly one such unit at the instant the reference is fixed, and (b) that *s* is being used at *t* to fix the reference of 'meter'. But knowledge of (b) depends essentially on sense experience, so it is obviously not a priori. Consequently the knowledge that *s* is one meter long at *t* is not a priori either.

Kripke's example never persuaded me, for the reason just given. Those who *are* persuaded by Kripke's example will find in my argument further, independent evidence for the possibility of contingent a priori knowledge.

p5, 'Contingent A Priori Knowledge', John Turri, *Phil. and Phenom. Research*

s is one metre long at *t*.

My take on this is that there is a conceptual connection between 'one metre' as defined, and the idea of the length of *s* at *t*. Now, as Kripke makes clear, this connection is not one of identity of sense; 'one metre' rigidly designates a length, whereas 'the length of *s* at *t*' does not.

Turri's picture seems to be that the only important connections between 'one metre' and 'the length of *s* at *t*' are: (i) actual sameness of reference and (ii) the historical fact that the reference of 'one metre' was fixed by

means of the phrase 'the length of s'.

I want to take a course midway between the sense-identification Kripke warned against, and Turri's separation. The simplest way to express my view is this: while 'one metre' does not mean 'the length of s at t', it is *part of the meaning* of 'one metre' that it refers to the actual length of s at t.

Let us say that two expressions have the same *two-part meaning* if they have the share such parts as just mentioned (definitional connections, roughly speaking), plus the same reference. (No presumption is being made that there are not other factors in meaning in general, by the way.)

This is still a little ambiguous, because being definitionally connected to the actual referent of some definite description is itself indefinite between, roughly, reference-only, sense-only and two-part connection. That is, here we have three ways of individuating definitional connections. Let us stipulate that we are talking roughly about definitional connections individuated in a broadly sense-only manner. (But here we want to give senses to proper names - something like individual concepts, internally individuated.)

Now, Kripke's point can be put this way: 's is one metre long at t' is apriori, in that any thought which is identical to it in terms of two-part meaning, will be true in virtue of meaning (definitional connection). How is this compatible with the contingency of s's being one metre long at t? Well, in any such world, there simply is no thought which is two-part identical to the thought in our world that s is one metre long at t. You can't have both the definitional connection *and* the identity of reference of 'one metre' in that world.

If one defined apriori to mean something like 'knowable no matter which world one is in', there would be no contingent apriori. (Coming to grips with the contingent apriori forces us to wise up about meaning or content - as it were, the fact that it doesn't float free of the world.)

'This house is its actual height' - this is contingent. There are possible worlds where it is false. But in those worlds, one can't have a thought identical in both sense and reference. So what does it mean to say that it is false at those worlds? Here, it is best to chuck world-talk as an explication of contingency, and say instead: it is contingent, because our system of representation permits us to represent an alternative state of affairs - albeit in the form of an essentially-counterfactual thought. For example, the thought of this house being half its actual height. Contingency is made within the system, but of course the system's adequacy is often an external matter. Contingency doesn't mean: cart all this apparatus off to another world, and the truth-value changes. It means: this apparatus is to be endorsed, and it can represent an alternative here.

Still, the only reason we said the thing was apriori is that in situations where it doesn't obtain, one can't have the relevant thought. One might worry that this gives one a kind of suspiciously cheap knowledge about 'which world one is in', or 'which states of affairs obtain'. As though one could *infer* something from one's ability to have such-and-such a thought.

* * *

'My boat is as long as it is' - contingent apriori. But what is it one can know apriori? The length of the boat? Then one could know the length of anything apriori. One could even avoid the sound of a tautology by defining a unit of measurement for each object, using that object itself. 'The length of my hat, I call 1h. Now, my hat is 1h long.'

How would it be if one could not see the object, and furthermore had no idea how long it is by any other units of measurement except for the specially defined one?

'Do you know what happened in the meeting?' 'Yes. I use the name 'M' for the exact pattern of events which in fact took place in the meeting. So I know what happened, namely M.'

What is the big deal about all of this? Can't we just shrug our shoulders and say 'That contingent apriori stuff again!?' Well, there are a couple of worries which come up:

It is disturbing to us that these silly procedures apparently satisfy our superficial criteria for 'saying how things are', and yet they plainly fail to say how things are. (For example, it upsets this "theory": when someone asks to know the length of something, a satisfactory answer consists in a well-defined expression which refers to that length.)

Similar puzzles for identity generally - knowing who someone is, etc.

We instinctively apply the picture of a gap being filled, rather than a connection being drawn.

Of course, there is nothing wrong with the idea that a gap is filled. Perhaps a safeguard for this picture is that the filler has to come from somewhere else, so to speak. And often not just anywhere, either.

Different uses of 'Who are you?' and 'Who is this person?' - to get a name when one can see the person, asking someone who can't be seen clearly...

There are extremely open-ended uses of this question. 'Who *are* you?', when you e.g. know the person's name and have seen them several times and spoken a lot - sometimes this is a rhetorical question, but sometimes it as it were says 'What am I missing? Tell me something about yourself which will make things intelligible here.'

* * *

What makes the contingent apriori puzzling? It is clear that we can construct essentially counterfactual thoughts about contingent matters, e.g. of a boat being longer than it actually is - so too can we construct essentially actual ones: the boat (if it exists) is as long as it is, S at time t is 1m long.

The puzzle is that it can look as though this allows us to partially locate ourself among all the possible worlds. Since we are apprehending the truth of something contingent, we are in some sense ruling out all the possible circumstances in which it is not so. It is that picture which plays havoc with us; one can't solve a murder by stipulating that 'N' designates the murderer, and then announcing that N did it.

This may make it look tempting to say something like: such propositions do not constitute substantial knowledge because they are sustained by the internal constitution of the host conceptual system. But this is inadequate, for necessary a posteriori propositions fall into that category, and are substantial.

We keep going for an over-simple, clear-cut characterisation. The last paragraph involves an attempt to replace the most naive conception, on which any answer to 'Who is the murderer?' consisting of an understood proper name which refers to the murderer is good enough.

There seems to be something in this: the contingent apriori proposition 'M is the murderer' does not constitute knowledge because it reports a connection which is made in grammar, in our conceptual system.

But there is a sense in which this holds for substantial identity statements as well. Thus it can come to seem like the previous thought was in fact worthless. (This is almost a picture of what has happened to logico-semantic philosophy, and Wittgenstein's insights, after Kripke.)

Reference and Representation

What determines reference?

In general, not descriptive or conceptual content associated with the name. Although the usual counterexamples to descriptivism do not apply to a range of names - for example, names of mathematical objects.

In general, not a causal connection between the object and the name. Both because the kind of connection in question is unclear, and because there is clearly no such thing in many cases - e.g. mathematical objects. But there is no clear counterexample in a range of cases: names of spatiotemporal individuals which someone in the linguistic community at some time has been perceptually acquainted with, for example. E.g. the baptism of a person, their name becoming widely known and passing into history.

What determines reference? - What gives this question its power?

The question seems to be: given that someone is referring to a particular object, what determines which particular object that is?

This is to be distinguished from: What is it which characterizes all cases of representation of, or reference to, an object? I.e. what do they all have in common which distinguishes them from other phenomena, apart from 'involving representation of an object'?

But it is close to: Given some particular object O, what characterizes all cases of representation of, or reference to, O?

The thing we just can't accept is that we have nothing definitive to say in answer to the question: what determines reference?

It is extremely remarkable and important that people have tried hard to say something definitive here. It not a matter of course that anyone is moved in this way by that question. (Many people innocent of philosophy would find the whole thing quite unnatural and meaningless.)

* * *

One thing I still have to talk about: the feeling that, when the investigation goes empirical and scientific, we have left the realm of essence, so to speak. But is that right? Could one not maintain that by increasing our knowledge of the contingent facts - how reference actually gets going - we may be better able to penetrate to the essence? I.e. to what it takes "in principle" for reference to get going?

Well, historical development itself cannot be part of the essence. It is clear enough that there is no logical impossibility in a whole lot of organisms being spontaneously generated out of nowhere and beginning to speak of things.

When we think of the representation of external, physical objects, some kind of correspondence or covariance theory seems attractive. One thinks of primitive representation (or proto-representation) which is of this kind.

However, the merest thought of arithmetic, and Wittgenstein's builders, at once makes all this look inessential.

There is no reason why one cannot erect a concept which applies, for example, only to the representation of physical phenomena. But if one is inclined to say that we refer also to abstracta, and to wonder how reference in general is possible, then an investigation of the first concept alone will never satisfy.

It is clear enough that the idea of referring, which is primitively connected with the ideas of looking-at and pointing - the picture of representation and represented - is widely useful, and widely used.

Now, it looks very much as though one might, having noticed this, ask for an explanation of why it is. I.e., what makes the picture of representation and represented apply in this wide range of cases. Or better: what is it about all these cases which makes the picture of representation and represented apply to them?

This traces the problem down quite deeply. At this point we see that none of the common 'theories of reference' is going to help us here.

And also, we start to see connections with other fundamental problems of philosophy.

For one wants to say: no real answer comes to this question. Once you are talking about a potential object of reference, you have already prejudged, so to speak, that the picture of representation and represented applies. Wherever we think of an object, there already we have the material for a higher order thought about our representation of that object.

Thus the problem of the essence of the representation of things leads, in a certain sense, to the problem of the essence of things.

This is instructive, because the problem of the essence of things is obviously no ordinary one! (Certainly no scientific one!)

To explain this tracing again: whenever we find ourselves in conditions where we can refer to an object, we thereby find ourselves in conditions where we can apply the picture of reference.

But, this is not the same as saying: wherever the conditions for objecthood are fulfilled, thereby are the conditions for a representation. For this is refuted by imagining a world in which there is just one object and nothing else.

This sort of distinction is important all through logical philosophy. It appears to be neglected, for instance, when people try to analyze existence and identity statements as designation and codesignation statements respectively.

Two senses of truth conditions, two conceptions of 'fulfillment worlds' for representations:

- (1) All conditions/worlds in which what the representation actually represents to be the case is the case.
- (2) The narrower set of all conditions/worlds in which the representation exists, represents what it actually represents, and where what it represents is the case.

What I was saying above now comes to: all the fulfillment worlds in the second sense for 'x exists' are automatically fulfillment worlds in the first sense for "'x' refers to x'.

Or: once 'x exists' is fulfilled in the second sense, "'x' refers to x' is fulfilled in the first.

And naturally, this holds not just for 'x exists' but for any proposition which contains 'x' in an extensional context.

By dwelling on this we can start to see that reference is no ordinary relation.

As long as one is prepared to ignore the distinction between (1) and (2), or to openly switch from one to the other mid-analysis, then one can analyze existence and identity both in terms of reference. (Heathcote.)

We might say: Reference is that relation which is presupposed when any relation is spoken of.

(This ignores some complications (e.g. fiction) but is far more appropriate than:

- (1) Existence is that property which is presupposed. and
- (2) Identity is that relation which is presupposed

because that distorts the logical form of existence and identity propositions. Reference is, so to speak, more properly a relation (i.e. more properly a property-or-relation, i.e. thing which can be ascribed to things).

What do I mean by 'more properly a relation' here?

Here is one way to explicate this. Imagine a form of representation where objects are represented with dots or boxes, which are then labelled with names (or otherwise made into unique elements). Properties can be indicated by labelled lines which have one point of contact with the object set, 2 place relations with labelled directed lines which have two points of contact, etc.

With such a technique, an existence property-representation and an identity relation-representation would serve no purpose. Every object would have exactly one point of contact with an existence line, and exactly two points of contact with a single reflexive identity line.

(This is a consequence of the way I am imagining this mode of representation to work. There are of course more sophisticated possibilities, where one uses dots and boxes at a higher semantic level, so to speak - almost as representations of ideas of objects rather than objects themselves. In *that* case, one can meaningfully use *something like* existence and identity lines - but then aren't they more properly seen as ascriptions of the property of designating something and the relation of codesignating respectively?)

Now, what about reference? Here we start to see that reference propositions can be sorted into two classes, two different logical forms. We might call these classes 'internal' and 'external', or 'semi-immanent' and 'non-immanent'. (This will become clearer in a moment.)

Consider what happens when we use this mode of representation and a new (extralinguistic) object comes to our attention (i.e. comes under our representation). We draw in a new box or dot (let us say "node" from now on).

But in so doing, we bring into existence a new object - one which stands in the representation relation to the original object. We can if we wish represent this fact now too. One way of going would be simply to draw in another node, which represents the last node drawn, and then to connect it to that node with a line indicating the representation relation. But of course this procedure can then be repeated for the new node. This procedure corresponds roughly to giving a name its own name in word language.

Another way of going, which corresponds roughly to quotation in word language, would be to introduce a kind of operator on points of contact. An unmarked point of contact is the ordinary case, a marked point of contact indicates that the contacted node is representing itself.

Consider the first technique, where we start at the bottom level, and then construct nodes to represent those nodes, etc., as required. This process thus "automatically generates" representation-relation ascribing propositions. These automatically generated ones are the internal, semi-immanent ones. All the others - i.e. when we talk about representors which aren't part of the system in use - are external, or non-immanent.

((Sketch of developments..)) Two kinds of criteria for non-immanent representation. The situation of the field-linguist. Direct comparison with our systems (looking at arithmetic talk for instance), vs. more complex coordination which somehow "involves the object" more.

There is a close connection here to the Twin-Earthable/Non-Twin-Earthable distinction.

Consider Idealism here. On Idealism, perhaps all non-immanent reference criteria can be reduced to system-coordination. A kind of collapse of Twin-Earthability.

For example: I see a man point at a rock and say 'N', and I form the hypothesis that 'N' is a representor (which is not part of my system) representing the rock I see in front of me. This is a paradigm case, it might seem, of acting on criteria for non-immanent reference which cannot be reduced to system coordination. I coordinate a part of his system directly with its object, not with a part of my system.

An Idealist may insist that this is not so. Rather, I am correlating my perceptual representation of the rock with something. But it would seem that if we are to be consistent, we can't really say that the perceptual representation of the rock is correlated by us with the man's representor 'N', since 'N' is not part of our system. We must say that we correlate our perceptual representation with **our representation of** the man's representor.

But still, this does not give the Idealist the distinction I want to have between external-object-involving verifications of non-immanent reference-propositions and verifications which involve "merely internal" comparison of systems. For in the internal case they **also** have to talk about **our representations of** some non-immanent representor.

Why is this important? **Not** "in case Idealism is true" - I wouldn't even talk of Idealism being true or false.

* * *

We must pay close attention to the language that we use to pose our problems over reference. 'How is reference determined?', for example.

We must interrogate and scrutinize the compulsion to think that we have a question in our grasp at all.

'x refers to y' is a flexible linguistic form, has a wide range of application. It invokes a picture or paradigm which we find it expedient to use widely. But if there is an analysis as Kripke speaks of in *Naming and Necessity*, it would be long and complicated. (So he says.) But how could such a set of clauses have exactly the same indefinite range of application as the original, simple form?

The notion of 'open texture'. Consider the fact that the concept of reference evolved. Now, when did the concept grow into its hypostasized analysis? The very last time it underwent change? And does everyone have exactly the same concept (use-profile)? Or would any such variations be mirrored in the application of the analysis?

The nature of the task or problem is the issue I need to press. I can't skip over that and go straight to showing it to be impossible of completion, as I used to want to do.

* * *

I have long had the feeling that the desire for a classical analysis of reference ('x refers to y') comes from within a way of looking at things which involves something like a "transcendental illusion".

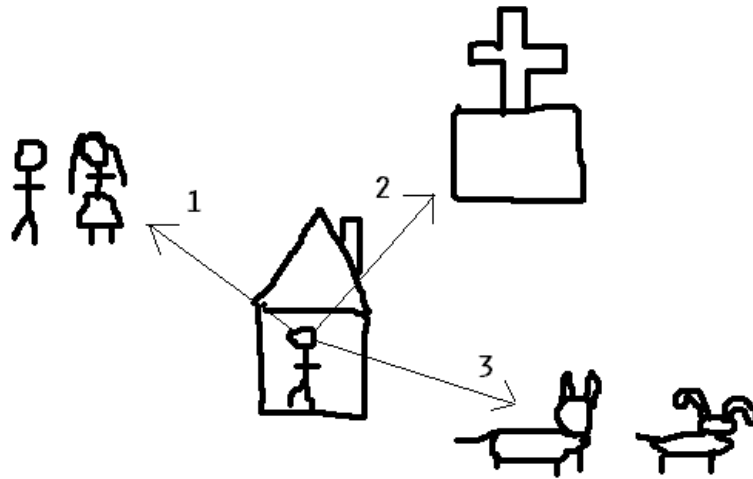
Roughly speaking: in real - non-transcendental-illusory - reference-talk, we always presuppose - internally, as it were - a connection between our concepts and their objects. Thus: what *makes* our idea of X into an idea *of X* we cannot so to speak answer to ourselves, from ourselves. But we can take up an external viewpoint: we can look at someone else's scheme and ask what makes *that* concept of theirs a concept of X. (And this will in one sense involve comparison between concept and concept, but in another sense between concept and object.)

Here there is the urge to say something like: 'Remember, you can't get outside your own mind'. But what the hell kind of reminder is that?! Who would ever think otherwise? (Satisfaction with such remarks belongs to a relatively low level of philosophizing.)

I can imagine someone proposing a 'two-component' theory of reference which makes use of a "disquotational base" together with "equivalence criteria". So when I say 'my concept of John is of John' this comes from the base, and so is similar to 'I have a concept of John, who exists'. But then when I talk of someone else's concept of John... well, I can do it stipulatively like that, but I can also say, for example, this name 'X' seems to be used for John. I.e. this word is used in the same way as my 'John'.

* * *

A man is in his house and thinks in succession of the baker's wife, the church, and a certain cat.



It is possible to be struck by the fact that this is the simplest sort of picture of what happened, and yet, in a sense, nothing of the sort happened; arrows didn't shoot out of the man's head. The puzzle is: How is this a picture of what happened? (Wittgenstein - the picture is clear, its application remains in the background.)

* * *

Alternative ways of posing the cardinal question in the "theory of naming":

(1) 'What is the meaning of a name?' - Unclear. Illicitly poses as a free-standing question which makes sense outside of a large, messed-up theoretical context.

(2) 'What kind of semantic values should be assigned to proper names in a semantic theory?' - A phrase like 'the semantics of proper names' might act as a kind of diffuser when talking to an outsider: (1) might seem silly to them, and they may quickly decide they know better than you and the theorists. But if you start bandying 'semantics' about, they might think 'OK, not my affair. Leave it to the experts.'

What exactly is a semantic theory? Must all semantic theories pursue the same kind of goal? What goal would that be? (To "explain linguistic behaviour"?, to "account for semantic intuitions"?)

It is not yet clear what business I have discussing this - whether I ought to be.

* * *

'How does language get to be about the world?' - If by 'language' is meant something which is of essence about the world, the question reduces to "How does language get to be?". And this is, in its way, a legitimate question, which may stimulate all manner of investigations, especially scientific investigations.

But this is not 'the problem of representation'. (Displaced wonder. There is no 'trick' to reference', etc.)

* * *

Perhaps the main thing Kripke did in "the theory of reference" was to show the inadequacy of descriptivism, conceived as a kind of *explanation* of individual reference, where the descriptive, reference-determining content is general-conceptual. (But it is easy to overlook how radically wrong it is -- this para belongs with the one about deferring the problem.)

And yet people speak of a theory - the New Theory, the Direct Reference Theory, the Causal Theory, as it might be.

"Direct Reference" - what does that mean? names aren't magic, front-row seat tickets to objects. But names are in a sense directly about particular individuals, in that they refer via concepts of particular individuals, unlike definite descriptions, which refer via general concepts. (But that way of putting the difference, though not incorrect, may mislead, for we also have this kind of situation with general concepts: 'round' e.g. is rigid, direct, and 'the shape of this model' is nonrigid, indirect.)

"Causal Theory" - well, this really doesn't follow at all from the antidescriptivism and the externalism. (BIVs.)

* * *

Kripke gives a picture suggestive of a kind of theory about reference, but it plainly won't apply to all cases where we talk of reference - say, to numbers. So if the 'picture' is supposed to be workable into a theory, as Kripke indicates might be possible and as others have attempted, how can the intended scope of this theory be specified without reducing it to triviality?

Perhaps this scope-specification cannot be independently made - but perhaps that is not any kind of objection, or justification of the label 'trivial'. Perhaps it's OK for the Kripkean theorist to say: my theory applies to the cases it applies to - and what makes it worthwhile is that it clearly applies in a lot of cases.

Kripke's sketch, as he admits, is no good as a non-circular analysis: it involves the notion of reference in the initial baptism stage, and also at each reference-passing, since he says that one of the conditions is that the receiver must intend to refer to the same thing (to rule out the case where someone thinks 'that would make a nice name for my dog', for example).

Mere recently we have Evans's full theory, based on the notion of the 'dominant causal source' of the utterance (or thought). But what are the criteria for dominance here?

* * *

EDIT

How do we manage to refer to things?

It is important to see that this is not like the case of a magic trick. There's no trick to representation.

We must forget answering it and speak straight to the spirit of the question, questioning it.

We could cynically view the desire of an analysis - a set of necessary and sufficient conditions couched in quite different concepts from the one to be analysed - as the desire to avoid having to deal with what bothers us here. But then we find that every sort of account which is proposed is either bothersome *as an account*, or clearly materially adequate but no improvement with respect to our initial problems, i.e. trivial.

'Reference' refers to reference. And '11' and 'the first prime greater than 8' refer to the same number.

There is a curious lack of discussion of such cases when, for example, Kripke-inspired causal theories are discussed. But the problem is serious: for it is no mere coincidence of language that we use 'reference' in both concrete and abstract cases.

Perhaps it comes down to this: the theory often applies - be happy with that. I.e. no non-circular specification of intended scope.

But then we get two equally urgent worries forking off. On the one prong, does such a theory *ever* work? I really have seen nothing which suggests so, and when I think about it I see that it really isn't something I would expect to be (conceptually) possible at all. On the other prong, what about all the other cases, then? Can such a limited theory be what we want? Are we not after, so to speak, the essence of *all* representation?

'A man is by himself, doing various things, and at 2:03 he thinks of Bertrand Russell. He is simply cleaning the house, or some such thing, and not speaking at all, and the thought is quite inconsequential, so much so that he is barely conscious of it. Nonetheless, it is a fact that at 2:03 this man thought of Bertrand Russell, and no one else.'

It is as though a kind of awe about human language, thought and civilization is being excessively localized, concentrated on the 'phenomenon of reference'.

* * *

In a paper called 'The Illusion of Semantic Reference', the author says, of the view that reference is 'primitive', 'unanalysable' - the view that there are no non-semantic necessary and sufficient conditions for reference: 'The problem with that theory is that if you say that, then if someone sincerely wants to know what reference is, you can't give an honest answer.'

This is an interesting statement, because that is obviously *not* the problem. No one sincerely wants to know what reference is (in the relevant sense) who is also familiar with the use of the term in philosophy. Which is to say: if someone sincerely wants to know, we *can* tell them, without giving a theory of reference.

* * *

It is curiously easy to put the *use* of the form 'x refers to y' to one side, as if it's a non-issue, or somehow obvious: "we use the 'x refers to y' form to state that particular connections hold between language and the world". It seems *perfectly* simple - a statement of a symbol-object mapping which either exists or does not. And of course that is a "correct" picture, a good picture. The complexity lies in the application.

This is connected with the tendency to view reference as a scientific problem. I.e. we all know what the phenomenon is, have seen it many times. The object now is to penetrate it and explain its working and origin at a deeper level. Like we might for lightning. But reference isn't like lightning (in the same sort of sense that the number four is not like a crayon).

On the one hand: there is a hell of a lot that could be said about this question, scientifically. But that doesn't mean we now have a 'theory-shaped hole' in our understanding, or that *that* is what is really bothering us. ('There is no trick to reference.')

"How the hell did you do that?!!" - the person won't say, or perhaps doesn't even know themselves, but a genius might be able to work it out. There are such things. Reference isn't among them.

* * *

We can become astonished that we are able to think of particular people even when they are far removed in space and time. That we are able to 'effect such a connection'. And as astonishment, there is nothing pathological in it. The pathology, I want to say, comes when one asks, as it were: 'How is this trick managed?' For it is no trick.

The thing I wish to combat is the idea that a simple but ingenious philosophical theory could 'explain representation'. And what of the semi-scientific writers who talk about representation and intentionality and wish to answer problems in this area with advances in neuroscience etc.? There, I think, something else is happening - something else is the primary thing - and it is no concern of mine. (The question-forms 'How is X possible?', 'How does X occur'" etc. admit of many different kinds of uses/answers. Consider: 'How does

a person move their arms?'. There is not just *one* thing which we could call 'how people move their arms'.)

In short, the problem of representation may serve as a stimulus to scientific investigations. But for the philosopher, this is by the by - the philosopher is not concerned to be thus stimulated, nor to interfere with science. The characteristic thing about the philosophical investigation here could be expressed by saying that our work revolves around, not the question 'What is the solution?', but rather 'What is the problem?'. ('We want an account' as 'Something positive had better come out of this!'.)

* * *

Following Kripke, Putnam, Evans and many others have said a lot of interesting things about linguistic practises - in a word, reference propagation and borrowing. It is not for me to get deep into the details of this. My whole problem is already present in the simplest, most *immediate* sorts of cases.

Here we see one way in which the problem of representation can be a stimulus; it can lead thinkers to investigate, loosely speaking, how some representation piggybacks on other.

I have to leave that to others and concentrate instead on the essence of the philosophical problem. And I think I have to be quite strict about this, or else what comes out will seem like an inadequate treatment of something else. (Something of this sort applies to much of what I'm doing here.)

For us, the problem is not a gap which needs filling, but more like a gap which must close up, by contracting to nothing. (My whole approach to philosophical problems could be summed up by saying that I am not interested in developing any kind of prosthetics.)

* * *

What I want to say about Evans's theory of reference is that it can only work by instituting a special understanding of the meaning of 'dominant causal source', bringing it very close, *conceptually*, to the concept of reference itself.

The way his view is commonly put is: the referent is the dominant causal source of the information associated with the name by the speaker.

So first thing is to see that it's not saying: the referent is the dominant causal source of the speaker's association. For then historical names would refer to history teachers and the like.

No, Evans - it would seem - wants to talk about the source of the information itself. (The original source, we might say.)

'Source of information' ambiguous. A source of existence, not a source of presence. A producer, not a conduit.

Now... what I think is that when we do enough disambiguation and really push this, we'll see that the information in question - its very identity - is not something which comes 'prior to' reference. After all, 'water is drinkable' arguably expresses different information for a Twin Earthian than for us, even if the representations are internally identical. Their projection isn't.

Quote from a biography (obituary) on Evans:

Information and Nonconceptual Content

The notion of information, as Evans used it, is not the notion of what a subject believes. Indeed, Evans suggested that we should take the notion of being in an information state as a primitive notion, not to be explained in terms of belief, judgement and reasons.

Evans 5.2. The informational system. Here he explains he is taking 'being in an informational state with

such-and-such content' as primitive.

But this makes it clear that the core problem of representation will be left untouched, whatever else Evans does manage to do. For we can get puzzled over how informational states can be about particular things, just as we can with thoughts and beliefs.

(One crucial difference here between beliefs and information states is that being in an information state with the content that X does not entail belief that X. For example, with the Muller-Lyer two line illusion, these will come apart. Also, following fictional stories.)

* * *

I must face the fact that the 'problem of representation' may stimulate someone to scientific achievement, under the confused impression that this will solve the philosophical problem. And I am trying to dissolve the problem - or rather, see how it comes about and get above it, without turning away from it. Thus it may seem that parts of what I am doing are inimical to scientific progress. To begin with, this is for the most part not true - the thing has been grossly overstated. But to dismiss it completely would be quite false of me. Why shouldn't there be a real tension here?

(I am reminded of certain success stories, for example in Hollywood. Matt Damon and Ben Affleck, or one of them, said that they only conceived and went ahead with pitching and making *Good Will Hunting* because they were too naive and ignorant to realize what a perilous thing they were embarking on. Thus, perhaps, if someone had 'educated' them about the industry earlier in their lives - genuinely informing them and increasing their understanding - *Good Will Hunting* would not have been made.)

One thing which might go some way to resolving the tension is to reflect that perhaps to a great extent we can identify different needs for different turns of mind.

* * *

One may think of a friend who is on the other side of the world (as I just did) and feel a kind of astonishment at the fact that this is possible. One is inclined to use phrases like 'definite connection' - as though it is a kind of surprise that they apply.

Do I want to criticize *this*? No - to criticize such a feeling of wonder as being, say, based on some kind of confusion, is the very last thing I would want to do.

And yet the movement from this to what I *do* want to criticize can be almost imperceptible. Rather than saying where it begins, I can say where it ends. And that is with the 'theory of reference' (or 'intentionality') as the theory of how language (or thought) 'gets to be about the world'. ('What is the *mechanism* of reference?' - SEP article on 'Reference'.) As though there were some trick which made it all possible.

And here again, I must be careful not to criticize the wrong thing. For reflection on questions of the forms 'How does this happen?', 'How is this possible?' etc., may serve as a kind of stimulus to genuine scientific investigations. It is not my business to get involved in any of that.

* * *

When one wants to understand lightning, the problem is, as it were, that it seems too singular. People looked for ways of seeing it as of a piece with other phenomena, and ways were found. With the problem of representation, it can seem to be almost the opposite. We look for a way of seeing representation as singular, but whenever we actually look at a case, the representation - the representing itself - always seems inextricably bound up with inessentials, as it were excessively of a piece with that which does not involve representation. We wanted to know what lightning had in common with other phenomena, whereas we want to know what distinguishes representation - what it has in common with itself only. But our picture, our concept, of representation does give us a unitary form - how isn't that all the singularity we could want?! It is as though we want to fill it out, replace it, but without losing any generality.

Here I wanted to speak of a 'yawning gap between picture and application'. But this is, in a way, a misleading form of expression, because a picture and an application are not similar, comparable things. To combat this particular pitfall, we might instead speak of 'a yawning gap between the

picture-plus-application, and other pictures-plus-application. (Frank Jackson's talk about 'location problems in metaphysics'.)

* * *

One might reason that the problem of representation can be traced to problems of mind. Linguistic reference requires intentionality - to give the words life. A linguistic term has reference in virtue of being used in a language, and this requires users who have intentional states. Thus in order to see what is really characteristic about reference or representation, we have to see what is really characteristic about thought. Thought is essential to representation - now all we have to do is say what is essential to thought. But wait - investigators of thought want to say that the representation is essential to thought! (Brentano.) So, it seems that we want to explain thought in terms of representation, and representation in terms of thought, but we also want to get outside this circle. Surely we can't have all three of these things! From my point of view it is tempting to say: OK, let's give up explaining thought in terms of representation. But isn't that simply because I am currently preoccupied with philosophical problems attaching most directly to representation, rather than to thought as such?

Perhaps I should instead opt for giving up trying to get out of the circle. Here my problems become especially acute. There is something deeply wrong with this triad. It paints us as cooped up, stuck - and we are stuck, but not in that way. It sounds stupid, but what we are really stuck on is the particular way of looking at the matter which make us feel stuck.

* * *

The problem of representational form vs. the problem of proxy representation, complex vs. simple

Complex: 'This diagram (or sentence) represents the state of the chessboard after the 10th move'

Simple: 'This pin represents Socrates', ' "Socrates" refers to Socrates '

This seems to be the distinction at play in the Tractatus and preceding work.

However, it is also used as a distinction between kinds of things. A kind of ontological, metaphysical distinction, apparently.

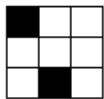
I don't go in for that. For me, the simple-complex distinction is a broadly logical distinction between kinds of representation.

I am tempted to say: to speak of a simple *object* is nonsense, unless that means simply that there is no possible complex-type representation which could be called a representation of that object. And that really does seem to be what it would come to.

Now the question: what aspects of the complex-problem come into the simple-problem? One might think: OK, so the object referred to by a simple sign, in basic empirical cases at least, is not fully determined by the structure of the sign system (Putnam's point), but surely there -are- structural conditions too?

A stand-alone representational system consisting of just one simple name - this seems like an impossibility, a fiction. To be sure, there might be a person who only uses one piece of language, a name, in -some sense-, but their doing this will be parasitic on their thinking and perceiving - on their natural representational apparatus. (By the time you get to something which could reasonably be called a name-object relation, one must have a good deal of other representational stuff *somewhere* on board.

* * *



Suppose objects like this, perhaps 3 dimensional, occurred in nature. Suppose that white squares correspond to opaque, cubic solids, black to an absence, a blank. Suppose these cubes are encased in glass (thus marking their boundaries). (Let us suppose that there is no gravity inside the cases, so that one could

consist for example of all gaps except one solid unit, 'levitating' inside the case, or held in by a wire frame - whatever.)

We might describe these objects using different representational systems. We have already seen two: the diagram, and my talking of solids and blanks at various locations.

Let us consider other possible verbal means of representation. At present, we are talking of a set of locations, at each of which is a solid or a blank. We might instead speak of a set of locations, at each of which is either a solid, or nothing (hence we say 'There is nothing there' rather than 'There is a blank there'). Or again, we might say that at every location there is a 'component', which may either possess solidity or transparency.

Using any these methods, we could fully describe the relevant properties of such objects. These descriptions, partial ones also, could be put into a correspondence with one another:

'5 locations contain solids, the other four contain blanks' = '5 locations have contents, the other four do not' = 'the components at 5 of the locations are solid, those at the remaining 4 transparent'.

"There are just 5 solids here' = 'There are just 5 things here' = 'Just 5 of the components here are solid'.

How about: 'There are just 5 solids here, but 9 things, just as there are 9 locations'? In the second mode, we can not say anything of this form: there, there are only 9 locations - there is no equinumerous set in the language game. This may look puzzling, as though the two modes were not equivalent after all, until we realize that in the first mode, it is a *tautology* that 'There are as many things as locations'. However, we could imagine extending the first mode of representation so that this is no longer tautologous, whereas there is no corresponding thing in the second mode. And so we might say that the first mode has a greater *potential* multiplicity, or points to a system with such multiplicity (whereas to similarly extend the second mode we should have to add a category, speak about another kind of object).

In the first mode, we can say 'There are blanks'. Nothing obviously corresponds to this in the second.

Or better: in the first mode: 'All structures which are not full of solids contain blanks.'

This sort of thing can create a serious puzzle for my way of thinking about modality. Expressed crudely: On the one hand, what is essential or necessary here seems to be, so to speak, the common factor between all these modes of representation. 'That in virtue of which they represent'. And from that point of view, a tautology from one mode which does not correspond to anything in another may for that reason look like something inessential, something which does not get hold of essence or necessity, but rather, reflects a peculiarity of the particular mode of representation which happened to be adopted, an artefact. And yet, on the field so to speak, we feel that these tautologies do have something to do with essence or necessity.

Could we say there are two kinds of essence which may be discussed? The essence of the phenomenon, and the essences of the various systems which represent it, picture it - and the latter must all subsume or mirror the former in some sense if the representational systems are to be adequate, but particular systems themselves will as it were generate further essences all their own.

This common factor, the means of representation common to all different means, apparently cannot be expressed. When we try to express it, we end up either with something (quasi-)tautologous, or something (quasi-) about signs. (Compare: 'John' refers to John, 'Mr. Smith' refers to John.) The point could be put this way: to describe the means of representation, one would have to describe both the phenomenon, the object, and the representational system, but to do the former one always has recourse to a particular representational system. But in what sense is there something we can't do, here?

Our problem can assume the form of a desire to avoid arbitrariness. 'What do all these representational systems have in common?' - in answer to this we can always pick one representational system, and use that to describe what they all represent. But then it would seem there are as many different answers as different

systems, or alternatively all are one in some sense, but then what is the criterion of identity? Can we just say 'they are all equivalent to one another'? But then comes the question: 'in virtue of what?'.

Could we have a model which, as it were, purely represents an object totally and without remainder - without anything inessential to the representation? A maximally 'efficient' model? But would that not just be the object itself?!

We have several equivalent modes of representation. Now, it seems like we can ask: in general, what properties must a representational system possess in order to be part of this group? But no answer comes, or at least no 'neutral' answer, no answer which does not simply employ one of the group and effectively say that all others must be equivalent with *this* one. We feel that there must be certain 'structural features' or the like in common, but we can never specify these without employing one particular system. And although employing one of these is always good enough in practise, provided it's an adequate and workable system, we feel as though it is never good enough for our particular philosophical purpose, because the system employed will always have inessential features, and we have not said which are these and which are the essential ones - or we have, but by means of another mode of representation, which itself will have inessential features.

I think this is what made the young Wittgenstein talk of that which cannot be expressed.

However, this will never be satisfactory, to simply lay it down that yes, there is that in virtue of which all these equivalent systems represent, but it cannot itself be represented.

This will never be satisfactory. Why not? Crudely speaking, we want to know what right we have to say such things. How is it that we can talk of 'something' which 'is the case', and yet which cannot be represented or spoken about? How do we know of its existence? 'It shows itself'. All these expressions certainly have their point, and it is most profound, but still: why is that? How do they work? We are using familiar linguistic forms in a highly singular way, and so one may want to ask: why *these* forms? But this will of course come to an end. (This brings us back around to the problem of the essence of representation, at a weirdly higher level perhaps: in virtue of what does 'the essence of representation is ineffable' represent?)

* * *

Could a causal theorist use set-theoretic, Frege-inspired reductions of abstracta as a way of bringing object-symbol causal considerations to bear on the representation of abstract objects? By holding, for example, that the number three just is the set of sets of three things? Well, one can't very well maintain that every time one represents the number three, this requires a causal connection with that set itself. How about maintaining that one must be connected with a part of that set? Well if a part must be a subset, maybe - but then there is the problem of being causally connected with a set **itself** at all. If a part may be an element, then the thing becomes trivial, for every object we are causally connected with is part of a set of three objects, and so we would fulfil the requirement of being causally connected to 'a part of three' just by being connected to any object whatever. Finally, the account just doesn't seem right - even if one grants that one can be causally connected to a set (or perhaps some kind of "fusion") of three objects. For pure arithmetical thoughts plainly do not involve that sort of connection, as might, for example, the thought that there are three houses in front of one. (One may handle the potential falsity of the three-house thought by saying that one has three representations of houses in one's picture of the scene - but this would sound ridiculous if the thought had been that there were 20,000 houses in the city, when looking down on it from a hill.)

* * *

How is it that x refers to y? This could mean:

How is it that x refers to **y** rather than something else?

How is it that x refers to **y**?, when y seems particularly inaccessible for some reason.

How is it that *x*, rather than some other expression, refers to y?

How is it that an expression like *x* refers to something like y?

How is it that anything refers to something like y?

How is it that anything refers to anything?

* * *

'N' refers to N.

When 'x refers to y' is not disquotational, it seems that for practical purposes it means x and 'y' codesignate, and then this can be understood as being verified by correlation of (aspects of) the role of two signs.

x refers to y iff x codesignates with 'y' iff x's referent = y.

' "N" refers to N in German too. '

Or when setting up a new language, as in formal Peano arithmetic:

' "0" refers to 0 '.

(The singularity of a truly disquotational reference-thought.) <-inadequate

What is bothering me is not fully clear yet, but it is something like this:

In philosophy, we think: 'N' refers to N. Then we think: how?

Now, that first thought is a thought to the effect that a certain symbol stands in the reference relation to a certain object. That much is clearly true to say.

Now, look at the thought as had by an English speaker, that 'Deutschland' refers to Germany. This too can be truly said to be a thought to the effect that a certain symbol stands in the reference relation to a certain object.

(Likewise the thought that 'John' refers to the man I met yesterday.)

But clearly the first, disquotational thought is a very different beast from the latter ones. What worries me is the effect of their assimilation under the rubric: reporting a reference relation between symbol and object.

Reference propositions are propositions which report such relations (connections), reference thoughts are thoughts that such relations hold.

So now we might say that the disquotational reference propositions and thoughts are a subset of all reference propositions and thoughts, characterized by the fact that the symbol which forms the subject of the thought is also used to represent the related object. And now we may ignore this subset and concentrate instead on the non-disquotational subset.

There are problems with this formulation, however. For example, suppose we are told that someone used a certain expression which itself is named E, but one does not know which expression E is. One may learn something about E, namely that it refers to Venus. Suppose E is actually 'Venus', the same word used in the same sort of language system. In that case, on my formulation above, the thought that E refers to Venus would be classed as disquotational, even when the thinker doesn't know that E is the name 'Venus'. (This is further testament to the variousness of the uses of 'x refers to y' and similar forms.)

((((don't lose the thread here - now, offer a fix, and then go on to the original point about a certain "account" being very natural for the quotational-but-not-disquotational subset - equivalence criteria. And perhaps nonquotational can be seen as *sort of* an existential generalization, or just a black-boxed, version of this - but that will then raise the issue of when what is in the black box is the symbol used to present its object - well, I guess it's not an issue: one just has to realize that here there is a limiting case, too - or, see it as disjunctive. Well, I'm getting ahead of myself since the restricted account's not meant to work anyway.)))

The difference between disquotational and other reference statements is reminiscent of the difference between trivial (repetitive) and informative identities. However, the problem is in a sense inverted: the naive relational view of identity statements makes all instances look trivial, including the nontrivial ones, whereas the naive relational view of reference statements makes all instances look nontrivial, including the trivial ones. (This use of 'trivial' may not be fully warranted, but suffices for making the point.) Identity statements are always necessary, reference statements always contingent - perhaps unless one is dealing with a non-Twin-Earthable symbol individuated at the level of full systematic role (or at the level of reference itself!).

It is instructive to compare ' "N" refers to N ' with 'The word "N" has reference, and it is used in the way it is used.'

(Or even just: 'The word "N" has reference, and it refers to the object which does in fact refer to'.)

What looks like (and in some sense is) a tautologous appendage changes the modal profile radically.

' "N" refers to N ', on the other hand, makes a show of referring to N with 'N' in saying that it is so referred to.

It is as though I am trying at all costs to avoid the picture of name and object. But I mustn't do that.

One points to an object - say a rock - and says 'This is called "N" '. Compare this with someone who knows the rock as 'M' being told 'What you call "M" is called "N" around here'. In the first case, the picture of representation and represented seems primary, but the picture representation and representation is applicable also (connection of a word with a cognitive-perceptual representation). In the second case, the picture of representation and representation (connection of two words) seems primary, but the picture of representation and represented is applicable also. And think how many cases are in between.

'London' refers to London. This fact could have an interesting historical explanation.

Our problem, however, comes in at the level of occurrences of symbols. It is clear that 'London' refers to London in ordinary usage. But what makes the latter occurrence of 'London' (the one not in quotes) one which refers to London? Not every occurrence of 'London' need refer to London. Suppose there is a Twin London. What differentiates London-referring 'London' occurrences from Twin-London-referring ones? That's one question. Also, suppose 'London' is used to refer to something quite different, say a person. What, then, differentiates a city-referring occurrence from a person-referring one? Suppose 'London' is used to refer to some fictional city. What differentiates the fictional uses from the normal? Suppose 'London' is used as a verb. What differentiates verb-usages from noun-usages? What differentiates occurrences which have some literal meaning from those which do not?

How do I know that this sentence actually refers to London, a city? How do I know that even this sentence refers to me, TH? The case of writing introduces complications which I wish to avoid at this moment. So suppose I just uttered both of those sentences, at the time of writing. How do I know that those utterances referred to a city, and to myself, respectively? How do I know they didn't actually refer to the whole of England, and to some rock which is called both 'me' and 'TH'?

Could something like this happen in principle? What would it look like? (Some story about a mental disturbance.)

It may seem like I am quite randomly getting into epistemology here, but I am not asking skeptical questions, or questions about the concept of knowledge. Here I want questions which admit of reasonably ready answers. And it is not hard to see how these might shed light on our concept of reference, of representation.

* * *

Daniel Dennett's notion of the intentional stance.

Can our question (or pseudoquestion, if it be such) about the essence of representation/intentionality then be put in this form?:

Under what conditions is it correct to take the intentional stance?

The question is whether this corresponds to the question of the essence of representation - not the question of the essence of representation of some particular object, rather than another.

The latter is the question about what determines reference, the former is the question about what determines whether reference is taking place. *This* question is untouched by problems with referent-determination theories, Quine's stuff about inscrutability and indeterminacy, etc.

Consider also: Under what conditions is it fruitful to take the intentional stance?

This is not the same question. For a start, it may overgenerate: it may be fruitful to take an intentional stance towards a system not because it is an intentional system, but because it is enough like one in certain respects, without being one. Where the intentional stance is treated instrumentally.

CLEAR UP THIS

(The real difference between an instrumentalist and a realist attitude toward a model has to do with the attitude one has. With the instrumentalist attitude, one doesn't get uptight about everything in the model being licenced or called for by the modelled. With the realist attitude, one does. Selecting and trying out models is like picking a suit off the rack - constructing a new model is like tailoring a suit. And of course, in many cases one will use a mixture of both activities: one will find something on the rack which needs some tailoring before it fits. Taking a realist attitude to the finished suit means treating any deviations from optimum fit for the particular wearer, any pockets they are left wanting, any unnecessary pockets, as defects in the suit. Corresponding to various meanings of 'instrumental', we have different possible attitudes to a garment: that had to off the rack suits, that had to a suit fashioned out of curtains by someone who is cut off from civilization, and that had to

* * *

I have the feeling that our puzzles about the essence of thought and representation are clouded by various things which are inessential for us here, and yet which are regarded as essential in the ordinary use of the relevant words.

For example, issues to do with consciousness and inner experience are, *prima facie*, no concern of ours.

It is also tempting to think: we must be prepared to abstract away from humanity - lots of features of human thought and representation are inessential.

However, we must proceed here with the utmost caution. For it is really not clear yet to what extent this is a good move.

With this in mind, perhaps a good frontier is this: to imagine and reflect on scenarios where robot computers (or things which are like them, in case that term implies something about intentional design or human origin) spontaneously generate.

One thing which strikes me immediately when I think of this sort of case is that 'representation' has a much better chance at a foothold here than does 'thinking'. 'Representation' feels literal, 'thinking' like a bit of a stretch.

Thought and intentions - these ideas point to a higher level of necessary complexity, so to speak, than does the idea of representation.

The idea of reference seems to be somewhere in between. Reference here is thought of as communicable representation, perhaps.

Distinctions in the vicinity:

"Private", internal representation vs. communicable.

Also: Natural signs vs. conventional.

An animal may have a representation of its environment encoded in its brain, but may not refer to anything using any signs.

Thus it seems like we might not want to take 'the intentional stance' to the spontaneously generated computers - especially if this means positing desires.

However, we may still want to take what could be called 'the representational stance'.

This stance, it would seem, is what is central to our problems in the philosophy of language and representation and their connections to modality.

(One doesn't need humanity, consciousness, and desires to have, e.g., something like necessary aposteriori propositions.)

Epistemic Modals

Indicative vs. subjunctive stance to epistemic possibility claims: what could be the case vs. what could have been the case. In the latter case, one is "referring" to a non-present (i.e. past or counterfactual) epistemic situation. In the former case, one's own is on the line.

Another distinction is that in John MacFarlane's article 'Epistemic Modals are Assessment-Sensitive' between context of utterance vs. context of assessment. This lines up, I think, with the distinction between being right to think that p could be the case, and being right in thinking that p could be the case. We can admit that someone is right to think, without being right in thinking, that p could be the case.

This most clearly applies to indicative epistemic modals, but it seems that it can also be applied to subjunctives. For example: John could have had anything in that box: when he showed us the contents

The language used for subjunctive epistemic modals, such as 'could have', seems more ambiguous than the indicative 'could be' construction. 'Could have' can be used to talk about what was (or would have been) reasonable to expect - this is the sort of case I have in mind when I talk about subjunctive epistemic modals. However, 'could have' is also used to talk about present epistemic modality, so to speak. For example: when wondering where John is, someone might say 'He could have gone to the pub', this is not about some non-present epistemic situation - it means the same as 'It could be the case that he has gone to the pub'.

In addition, it is used in quite other ways: it is used as a way of putting 'can' into the past tense: someone says they can do something. This is a claim about potential, ability, *not* an indicative epistemic modal claim to the effect that it could be the case that they will do that thing. Yet, for the past tense, the overworked 'could have' is the natural thing to use.

'Could have' is also used for grammatical, conceptual, and/or metaphysical modality.

'Could have' is thus used in at least these four discernable ways: (i) talking about a non-present - past or counterfactual - epistemic situation (the core subjunctive use), (ii) talking about what could actually be the case about the past, (iii) talking about lapsed potentials/abilities, and (iv) making conceptual or metaphysical modal claims.

With the core cases of subjunctive epistemic modals, there is need for a distinction (different from MacFarlane's) between context of *subject-matter*, so to speak, and context of assessment. E.g. 'White could have easily won at that stage' - here there is no prior claim which is now being assessed - the claim is both made and assessed now, but one considers the epistemic situation at an earlier time ('that stage') when assessing the claim.

Indicative stance: here and now is what matters. So - upon learning more - what one said earlier can be said to have been wrong, even if one was right to say it. So one wasn't making a claim about a past time and place. The claim itself is dynamic, one might say.

Subjunctive stance: what was reasonable to believe given some cognitive situation.

The indicative stance corresponds to the assessment of 'right in thinking' claims, the epistemic context of assessment is the live thing. The subjunctive corresponds to the assessment of 'right to think' claims where the context of utterance is the thing. My talk of context of subject matter is just a way of talking about, if it had been thought in such a context, whether it would have been right to think.

When considering our own present thoughts and claims, it might seem like these two line up. The question 'Am I right in thinking this?' and 'Am I right to think this?' become effectively identical. It can look like, in this special case, there is no possibility of two different stances. This is perhaps similar to the situation with 'If x is the case, y is the case' and 'If x were the case, y would be the case' when one is keeping an open mind about whether x is the case.

But what about those who know better than us? Obviously, this is a possibility we allow for. Thus it might seem that one can be quite sure one is right to think some epistemic modal, but then the question of whether one is right in thinking it might seem to remain open. Thus we must reconsider the previous thought.

With those who know less (including our earlier selves), we can take two assessment-type viewpoints on their epistemic modal claims: original context relative (right to think) and assessment relative (right in thinking).

But with those who may know more, it is very different. In this connection calling the indicative, 'right in thinking' sort of evaluation 'assessment relative' may mislead. The point is: with such evaluation, one

assesses against the best knowledge-base (belief-base, more accurately) one can muster - and the claim retains its identity when this live knowledge-base alters.

* * *

One thing seems clear: we can get some grip on the conditions in which it is right to think that p could be the case. The only trouble is that this doesn't entail that what one thinks is right. This seems to remain open for as long as one lacks definite knowledge one way or the other.

The puzzle is: under what conditions is 'p could be the case' is actually right? (Seeing why a puzzle is insoluble in one sense is a solution in another sense.)

Someone who we think knows more than us says 'It couldn't be that p' - whereas for all we knew, it could have been. We might believe them, and move to that view ourselves, or we might say: 'They might be right about that, but they might not'.

This gives a hint of how complex things are here - we might quite firmly think that it could be that p, and yet, if someone else says it couldn't be, we can consistently say they might be right.

Another puzzling fact: when considering whether something could (might) be the case, we sometimes say we don't know. As though we're saying: it could be that it could be, but it could also be that it could not be - where 'could's 1 and 3 are tied closer to the speaker's present state than 2 and 4.

A lot of our talk in this region seems to make reference to something like: 'the cognitive situation we ought to get into' - and this needn't have the flavour of some far-off ideal, but can often be quite easily attainable.

I say 'present state' because 'present knowledge' is more indefinite. (Consider 'I think I *might* know this' - in one sense, the whole point is that present knowledge may indeed cover the point at issue, while in another more immediate sense of 'present knowledge', it avowedly does not.)

'I used to think it could not be the case that p, but now believe it could be the case - my earlier judgement rested on grounds I now take to be false.'

Here, the source of past incorrectness is that what was taken as known was not in fact true.

Compare: 'I used to think it could be the case that p, but now I know that q, and therefore that p could not be the case.'

Here the source of past incorrectness is that I didn't know enough.

Two poles of description: cases where I've got a bad apple in the system, and cases where I lack some good apple. But there is also the case where what holds me back is naturally thought of as a misunderstanding, or a lack of understanding, about what apples I have in my system.

Sometimes it is clear that one is dealing with the first sort of thing: lack of good apples or presence of bad. I.e. in cases where one just lacks the relevant empirical information - so that no amount of reflection would help, and in cases where one has a well-founded but false empirical belief - so that no amount of reflection would help.

But this becomes indefinite sometimes - for example when learning mathematics.

'It could be that at the meeting, all 9 people shook exactly 5 others' hands'

This could be refuted both empirically - by learning how the meeting actually went - or mathematically (since graph theory shows the thing to be impossible). This example is more interesting than one involving, say, fairly complicated arithmetic, because someone unacquainted with graph theory and similar sorts of

mathematics may have no inkling that the thing is mathematically impossible; a quite intelligent person in our culture could easily fail to see a potential for 'apriori defeat' here.

Contemporary discussions of 'modal error' could probably benefit from consideration of such cases. They are quite different from the Kripkean necessary aposteriori: there, one doesn't know something - and this thing is necessary. In this sort of case one can know that the 9 people didn't shake hands with exactly 5 others, but what we don't know is that it's impossible.

Question: can one give in-advance criteria for whether a proposition which may look empirical might be refuted (or verified) purely mathematically?

It is clear that, without knowing all current mathematics, let alone all possible mathematics, we *can* safely rule out certain kinds of propositions ever being refuted or proved purely mathematically. For example, simple statements about the location of a single body.

One might look for necessary and sufficient conditions for being undetermined apriori. Failing that, one might try to give more and more sufficient conditions, perhaps yielding a longish disjunction, widening out a "securely empirical" realm.

* * *

My task with respect to epistemic modal discourse is to adequately describe its working, in a way which will throw into relief both the important similarities and differences between it and other things we call statements or propositions. (This also holds for many other things.)

It is easy to overestimate the similarity: hence truth-conditional analyses involving context parameters. It is also easy to overreact when this fails. 'These are not truth-apt propositions, they don't have truth-conditions' - it needs to be seen that this doesn't say anything very determinate - rather, it is more instructive to look at such philosophical utterances as an attempt to regiment the grammar of 'proposition', 'true', 'truth-condition', etc. (This point is crucial for "metaethics".)

* * *

Epistemic modality as pertaining to truth-value candidature.

- (1) Possible - candidate for truth
- (2) Not-necessary - candidate for falsity
- (3) Necessary - not a candidate for falsity
- (4) Impossible - not a candidate for truth

Regarding (2): think of the form of response 'Not necessarily; ...'. (Note that we don't use the word 'contingent' here because that presupposes truth.)

Ordinary 'could be's can be looked at this way. ('Dynamic epistemic modals' - to signify that they aren't assessed against any particular knowledge base.) Also, 'for all I know', 'for all we know', 'for all that's written here' (knowledge-base relative); statements about apriori knowability can be seen as knowledge-base-relative, the knowledge-base being what can be known apriori.

But - apart from 'relativity' problems with the knowledge-base approach (i.e. with dynamic epistemic modals), there are other issues here.

For one is tempted to say: 'It could be that p, for all I know' is true iff p is not ruled out by the subject's knowledge - or, iff p is compatible with said knowledge. But if the criteria for being ruled out, and being compatible, are purely logical, this will not work in cases where the matter for speculation is whether or not something is implied or not by some knowledge-base.

Nonetheless, some cases do seem to conform to the analysis given. For others, we might talk about

robustness to some fluid, contextually set amount and type of reflection. For other still, perhaps we could simply say that 'It could be that p, for all I know' is true iff the subject doesn't know either p or $\sim$ p.

Now, despite all this, when we leave such rigidly knowledge-base relative cases, we must also leave truth-conditional analysis behind, despite the similarity of ordinary epistemic modals to the intermediate, contextual-robustness case discussed above. Or so it seems to me. But this does not mean giving up on 'truth aptness' or 'cognitivism' about ordinary epistemic modal statements.

I think the way forward here is to leave the issue of truth-conditions to one side for now. There is a good deal, I suspect, which can be said about the conditions under which one is right to think and to say that something could, or couldn't be, the case. Also, about what happens in conversations when one person tries to get another to assent to such a statement, etc.

This is the primary thing - to look closer at these statements on their own terms. In the course of this, we will be able to see better how 'true', 'false', 'right' and 'wrong' fit in here.

We see immediately that there is something peculiar about epistemic modal statements, and we get tempted to mark the difference in an overly crude way: e.g. by saying that they aren't really propositions, or aren't really capable of truth, or are incomplete propositions. Or one gets a little more sophisticated and says: well, you can call them propositions, and say they're true, but they're not descriptive - they don't describe anything. But what about a case where someone says 'Maybe ...' where the blank is filled in with some long and richly detailed description of a scenario? Such a proposition has descriptive content. Perhaps the thing is that the content is not simply put forth as a description of how things are, but rather of how they (actually) might be. A tentative putting-forth of a description - 'we might need this'. So now this angle on the peculiarity of epistemic modals seems weak as well: we seem to be reduced to saying that they don't describe actuality but possible actuality - what might be. And that is trivial. Or we are just saying: we ought to distinguish this sort of proposition - it has its own characteristics. And that is not unwise, but the original aim of the non-cognitivist and the non-descriptivist (quasi-realist) was to say more than that. To get to some real (instructive) differentia, we have to look closer. (What the quasi-realist treats as the argument for their view is actually more like an explication of the content of their view. Remove the theoretical gloss, and all the important data remains - and in plainer sight.)

The weird debate between Blackburn and critics about whether language 'has a downtown'. The hard thing here is to let go of the feeling that anything in particular is at stake. In a sense, one must be more *direct* here - and talk about the facts which have impressed one. (Of course, the quasi-realist might say that this is less direct, is jumping around the issue, etc.)

The critic of the quasi-realist rightly takes issue with the *slant* which quasi-realism puts on the matter. To argue about whether there is a downtown or not, or whether there is a primary region, is not to the point. It would be better for the critic to go along with the talk of a downtown, a primary region, for the sake of argument, and then to show how little it buys one.

'These are the *real* propositions' - there is quite clearly something *evaluative* in this. ('This is what it's all about'.) To be sure, one can *defend* this with a series of in-depth logical observations, but this does not make the object of defence into a theoretical proposition, a scientific proposition. One might defend an aesthetic judgement about a piece of music with a series of in-depth musical observations, too.

(Goethe: don't look for something behind the phenomena - they themselves are the theory.)

* * *

This would be a natural thing to say: 'I'm not interested in what could be the case, I'm interested in what *is* the case.'

This, much less so: 'I'm not interested in what is the case, I'm interested in what *could be* the case.'

Interest in what could be the case presupposes interest in what is the case.

Of course, we can take indicative epistemic modals and preface them with 'It is the case that'.

'It is the case that N.N. could be coming.'

But 'is the case' in the natural utterance above, means what is the case *as opposed to* what *could be* the case. That part of what is the case (in our view) which does not concern epistemic modality, one might say.

This way of speaking might make it look like 'epistemic modality' denotes some subject matter which we speak of. You can say that if you like, but it is also naturally seen as - roughly speaking - a form of thought, part of the technology of thought.

* * *

If someone thinks that it could be that p, and you think you know more about the matter, you might think otherwise while granting that they are right to think it, given their situation. On the other hand, if someone makes an epistemic modal claim and you think *they* know more about the matter, there is obviously no space for thinking they are right to think it in their situation, while having a different view yourself. (This is trivial.)

If you judge that it could be that p, in one sense you may be able to rest perfectly content with this judgement - as it applies to your present situation, so to speak - but at the same time you might endeavour to find out more, and in the course of this you may confirm or refute the judgement.

One can be right to think, but wrong in thinking, that it could be that p.

Secondly, it seems that one can be wrong (unreasonable) to think that it couldn't be that p, even while being right in thinking it.

(Can one also be right to think, but wrong in thinking that it *couldn't* be that p? And can one be wrong to think, but right in thinking that it *could* be that p? These cases seem more outlying, but we can imagine things here.)

The general version of the first point is that it can be reasonable to make certain incorrect judgements in certain circumstances, the second is that it can be unreasonable to make certain correct judgements in certain circumstances. But the foundation of this point in the present instance, so to speak, is unclear - for it seems that there is no kind of correctness with epistemic modals beyond something like reasonable judgement given circumstances. It's hard to see what else could possibly come into it.

The *flexibility* of the technique of epistemic modal language is what we are unprepared for.

Think of what sort of things make the difference between whether someone thinks (says) 'It could be that p', versus 'I don't know whether p could be true'.

According to general ideas, derived from other parts of language, it might seem like these two thoughts or assertions must be utterly different. We are looking at the difference between 'q' and 'I don't know whether q'. (In this case 'q' is 'it could be that p'.) And yet, in an epistemic context such as this, these really aren't so different as the two schemas make it look.

(This reminds me of something I heard just earlier today, in ordinary conversation: 'Yes, that's probably quite likely'. In both this and case just under discussion, it can look like the "doubling up" (loosely speaking) has mainly to do with effecting something quantitative - force, strength, something like that. But it can also look like the two "operators" are not really the same: that one has to do with the present cognitive situation, and that the other deals with a better one.

But no, I want to say: this just comes from the nature of what you're iterating. This point must be made clearer:

'It could be that p ', by itself, says (as it were) that coming to hold ' p ' might be an epistemic improvement.

'Maybe it could be that p ', on a unified understanding of the two operators, says that coming to hold that 'coming to hold that ' p ' might be an epistemic improvement' might be an epistemic improvement.

This makes it clear that the fact that the second operator 'has to do with' a better cognitive situation than the first is not to be taken to suggest that the two operators are different operators. The first acts on the second.

(When I say that the operators say such-and-such, I'm not literally talking about literal meaning, but proposing analogies, or alluding to an area where two forms run close together.)

Presumably we never have thoughts involving three such operators. Assuming that is right - is it just an accident? Might more complex minds make good use of such a construction?

It probably can happen, in conversation, that three or more the same operator like this does build up, but I suspect that it's never significant for us beyond two. There is a subtle but still quite concrete difference between one and two operators, but no perceptible "real-life" difference between two and three.

One thing now would be to describe in more detail how the difference between one and two epistemic possibility operators can be important.

It is as though epistemic possibility claims get more and more definitive - more and more warranted - the further the relevant inquiry, discussion, reasoning or experiences goes, with confirmation or disconfirmation at the limit.

The puzzle, then, seems to be this: it looks in many ways as though we are approaching a limit - and it would seem that the limit is total knowledge. But epistemic modals have no meaning (as I feel like putting it) in the idealized context of total knowledge - there, epistemic possibility just collapses into truth, impossibility into falsity. (More accurately: 'it could be that p ' collapses into ' p ', 'it could not be that p ' collapses into ' $\sim p$ '.)

Thus it would seem that, if you want real truth conditions for epistemic modals such that, if these conditions are fulfilled, the epistemic modal in question won't ever let you down in the light of better knowledge, then you *must* say that 'it could be that p ' is true iff ' p ' is true, 'it could not be that p ' is true iff ' p ' is false.

This point, as put, is vulnerable to the idea of essentially unknowable truths. But the point can be recast in terms of the best possible knowledge state, rather than directly in terms of reality. Thus it seems that the "ultimate" truth condition for 'it could be that p ' must be 'it is not knowable that $\sim p$ ', and for 'it could not be that p ', 'it is knowable that $\sim p$ '.

Our problem can be seen, as can many philosophical problems, as a need to steer between two extremes. In this case, two extremes about the conditions under which 'It could be that p ' is correct (in some sense).

The thing to look at, presumably, is how epistemic modals can fail to be correct (in some or various senses) - and how such cases differ from the others. For plainly some epistemic modal claims are more warranted than others.

I believe one thing which makes this area difficult is that there are two factors in the assessment of 'It could be that p ' - epistemic warrant, and also whether or not p . Neither is enough by itself:

Robustness (from Frank Jackson's theory of indicative conditionals) seems to be a fruitful notion here.

When 'p' is the case and knowable, 'It could be that p' will thereby be ultimately maximally robust - although, of course, one may get into an unfortunate, misleading epistemic situation in which one gets more and more evidence *against* p.

When 'p' is not the case and this is knowable, 'It could be that p' might still be quite robust in certain circumstances, enough for some kind of correctness or truth.

On the other hand, it can happen that 'p' is not the case and 'It could be that p' is insufficiently robust for the claim to get a positive assessment. If a moment of the slightest reflection reveals that $\sim p$, then it is wrong to say or think that 'It could be that p'. (However, there can be uses which are not even meant to be *this* robust. 'Maybe, let me think...' (3 seconds pass) '.. yes.')

Now: can it happen that 'p' is the case, but 'It could be the case' is *not* robust (in a non-ultimate sense) enough to be correct? It would seem so. ((This point may need a little expanding in order to be understood.))

The original point was: the relevant knowledge state must correspond with reality. This can be modified to: the relevant knowledge state must tally with the ideal knowledge state.

* * *

'Could be' assertions have a different kind of role, depending on whether one takes one's audience to know more or less, and also depending on levels of mutual trust.

'Give me your personal details'

'No! You could be an axe murderer!'

'Could be's' in internal accounting, vs. conversation.

Thinking vs. saying.

The 'you could be an axe murderer' case above has a special flavour. It is common ground, in this exchange, that the hearer probably knows more about the point at issue. Other 'could be' assertions, on the other hand, are supposed to be more *directly* accepted by the hearer.

In the axe-murderer case, one can add 'for all I know', without affecting the practical sense.

* * *

For representing belief states, states of cognitive systems:

- (1) For each proposition, a sorting into BELIEVE, DISBELIEVE and AGNOSTIC.
- (2) For each proposition, a credence value.

There are various ways these two modes of representation could be combined. (Connections here with the 'theory of vagueness'.)

One obvious way: Start with (2), then set two thresholds.

Also, one could of course have a series of different thresholds corresponding to different standards.

Here, perhaps, is a lens through which to look at epistemic modals:

Credence reports, credence proposals, credence reports plus proposal.

Reports: 'could be, not sure', 'you could be an axe murderer'. Proposals: more normative force: you should be like this too.

(Intermediate cases: games where someone withholds knowledge. 'It may, it may not'...(The coy 'maybe'.)

One reason this lens is potentially illuminating (mixed metaphors!) is that we can *also* see regular contingent empirical descriptions/assertions in this light (belief-report plus proposal) - thus putting them side-by-side.

(Whereas it is more common to focus on the deep differences, to be sure, but only the superficial similarities. Wit PsychII 56 (unrel matter): 'The basis for comparing them is hard to see; not their difference.'. --- also the question 'How are all these things *propositions*?')

The challenge is to strike the right balance (in our overview of kinds of propositions) between similarities and differences (their emphasis, their placement).

BEL -----threshold-----AGNOST-----threshold----- DISBEL

Assertion: 'Put your credence in this to the left of the first threshold.'

Denial: 'Put your credence in this to the right of the second threshold.'

Assertion of epistemic possibility: 'Put your credence in this to the left of the second threshold.'

Denial of epistemic possibility (assertion of impossibility): 'Put your credence in this to

Could be false: 'Credence to the right of first threshold'.

Could be either way: 'Put your credence between the two thresholds.'

Thus, we have ways of selecting one region, or ruling one out.

We don't have an analogous way to highlight belief and disbelief, ruling out agnosticism.

We can highlight as 'go-zones': the left (assertion), left and middle (could be true), middle (could be either way), middle and right (could be false) and right (denial).

Now, what about assertion of (epistemic) necessity and impossibility? Like plain assertion and denial, they allow just the left and just the right respectively.

So: these considerations clarify an awful lot, but still leave us in the dark w.r.t. the question of how these two pairs (epistemic *necessity* and *impossibility* assertions versus plain assertions and denials) relate to one another, what the modals are for, etc.

It is easy to see why we might use 'could be' as opposed to 'is'. We might want to leave it open whether it is or not. But no parallel explanation is available for why we would use 'couldn't be' as opposed to 'isn't'. (Cf. 'Has to be' as opposed to 'is'.)

The difference between 'It could be that p' and 'It could be that ~p' isn't all that great, except in the limiting case where the issue of p is actually settled. Two kinds of cases: (1) speculation toward the truth, where one tends to embed the more probable proposition, (2) cautionary uses, where one tends to embed the less probable.

* * *

We need to investigate epistemic modals, but we must not expect it to be a bit of piecemeal work. We must be ready to let the investigation - its results - operate on us; on our conceptions of language and of a proposition. And *this* is why we must avoid trying to give an account of a certain sort - roughly, a sort which shows a way of bringing epistemic modals under a more familiar propositional umbrella. We must extend our concepts of language and proposition so that epistemic modals simply have their place, which is given partly by analogies or similarities with other forms.

//////////All this may be a bit too wishy-washy, in one respect. Namely, that it may make it look as though the lessons of epistemic modals are rather ineffable; we can't sum them up, we can only look at the facts and 'let them operate' on us. Well, I don't know about this. In effect, I am still trying to sum them up. (This comes naturally, whether I like it or not.) I think a good deal may be accomplished here. In particular, a kind of tabulation of the similarities to simple descriptive contingent empirical propositions, as well as the differences. Then one could perhaps say a lot about our troubles, by saying things like: we are impressed by *this* similarity, and it makes *this* difference look out of place, since we are used to *that* sort of case.

Also, the comparison with what I called 'simple descriptive contingent empirical' propositions is not the only fruitful one which can be made. Being specific in that way puts more meat on the bones of the investigation, so to speak. The broader category to look at would be something like 'those propositions for which the notion of objectivity seems as straightforward as it will ever be'. Arithmetical propositions get straight in. What about necessities which are not mathematical or logical?

Let us try for a preliminary capturing, in the most basic terms, of what is bothersome about epistemic modals.

On the one hand, it in some ways looks as though 'It could be that p' expresses a different proposition depending on who utters it. (This shows up in the way we are ready to say that a presumed less-knower is right to think 'It could be that p', even if we believe that it is not the case that p.)

On the other hand, it in some ways looks as though 'It could be that p' expresses the same proposition for everyone. (This shows up in the way we are ready to say that, even if a presumed less-knower is right in some sense to think 'It could be that p', they are not **right in thinking** this, inasmuch it "couldn't really be" that p.)

There is a characteristic phenomenology of, as it were, getting serious; asking if something could *really* be the case. It seems clear to me that the characteristic effect of such a question is similar to saying something like 'Let us to our utmost to go by our best lights, here', or even 'Let us see if we shouldn't throw this away'. (Popper on falsificationism.) *This* fact has something to teach us with respect to the problem of epistemic modals. (We are emphatically not, for instance, trying to get down to the business of what an ideal knower would think; in many cases it is taken for granted that we are very far from this in the relevant respect.)

This is connected, in a way which I want to get clearer about, with the fact that ordinary speakers very often want to use two modal expressions - the second being an "extra buffer", as I feel like saying.

Also, the fact that when asked if something could be the case, we often want to say 'I don't know', belongs here.

'Could this be so?' - in some cases one might confidently say 'Yes'.

Think of the case where one knows that one lacks a piece of knowledge which *may* decide the first-order matter in question in the negative, but which could not confirm it. In such a case (and others, too), when asked if it could be so, it would be very natural to say 'I don't know - because I don't know whether ...' ('...' indicating the object of ignorance). 'It might be possible'. (In the case where we know we lack a piece of upcoming knowledge which is sure to decide the matter, however, we would naturally say 'Yes, it could be. We will see when ...'.)

'Could he be at the beach?', when looking for someone. Think of what it would mean here to confidently answer in the affirmative. It would mean something along the lines of that there is a pretty good chance he is at the beach. Consider on the other hand what it would mean to answer 'Maybe' in an uncertain tone of voice.

* * *

UPTOHERE

Further things to work in: accessibility. Some knowledge seems totally inaccessible to some agents, and then there are lesser and lesser degrees. Different standards - something like Lewisian context at work here. (There seems to be a connection here to the issue of responsibility and freedom, failure by omission - not doing what is in one's power.)

* * *

A 'could be' belief or assertion seems more defeasible, less conclusive, than a 'couldn't be'. 'Couldn't be' implies 'isn't'.

What is it for a 'could be' to be conclusive? Apart from the trivial limiting case, where we actually

By 'defeasible' and 'inconclusive' I don't *just* mean revisable in the way that lots of propositions happen to be. With 'Could be' propositions, inconclusiveness is a (or *the*) crucial part of the language game.

The confusing thing is that one must *also* make room for (and sense of) different levels of surety with respect to 'Could be' propositions themselves.

When, apart from the limiting case, is a 'Could be' most conclusive? The answer seems to be: in cases where we know we'll never know.

What could it mean for an epistemic modal 'It could be that p' to be conclusively confirmed, once and for all? Possible cases:

- Where 'p' itself is conclusively confirmed.
- Where it is conclusively confirmed that knowledge as to whether 'p' is absolutely unavailable.

The fact is, however, that we - as users of epistemic modals - are not primarily interested in these sorts of confirmation.

What about "weaker" kinds of confirmation? Some of them can look conclusive enough in a way, too. I am thinking here of the way we use 'It could have been that p' in an epistemic sense, when we now know that it is not the case that p. To say roughly that it would have been reasonable to be open to p. Here it looks like we are agreeing with the epistemic modal as it stood in the past. However, this would be a very weird thing to say: A: 'What I'm saying is: it could be that p', B: 'No, it couldn't, because (...)'. Granted, what you're saying now *used to be* true.'

So, we don't exactly stand by these old judgements. We revise them, and then use a different but closely related form to mark the old situation. So, this 'could have been' stuff is perhaps sort of a red herring in this connection, namely the original question in the last paragraph, about "weaker" kinds of confirmation than limiting case and know-you'll-never-know.

There is a kind of provisional confirmation which is secure, in that the need for the provision is clear. For instance, cases where you know you can't know now, but will be able to know later. Here you might be quite confident about a 'Could be', even though it is clear that this may have to be revised once the foreseen knowledge-access comes about. But one might *also* play the game another way here, if asked whether it could be. In short, two ways of answering the question, in such a case, of whether it could be that p:

- Yes it could, but we have no way of knowing until ...
- We won't know until ...

The second would be more natural if the foreseen bit of knowledge-access may not get one all the way, but may rule out. (I think I was discussing this in ep pos 5.txt last night.)

A clear sort of setting for such a case: some detectives think that a gun found near the scene of a murder could be the murder weapon. They send it for forensic examination, knowing full well that the technicians cannot confirm that it is the murder weapon, but that they can rule this out. The detective think that

perhaps it could be the murder weapon, and they use forensics to 'find out' whether it really could be or not.

When we revise a could-be, two things might happen: we might look on our previous judgement as something we ought not have done, or we might look on our former selves as having been right to think it, even if not right in thinking it.

Now, one thing I think has to be said, is that when one asserts a could-be, one is *not* saying that revision will not have to occur. One is not committing one's self to that. However, one is "asserting", in effect, that even if the thing does get revised, this revision will be of the second kind, where we quasi-stand by our judgement (replacing it perhaps with a 'reasonableness' 'could have been').

This seems to be of crucial importance. I just need to find a 'bright' place to put it.

In place of truth-conditions here are the conditions under which one is right to say or think an epistemic modal. And that of course depends on how far one intends it to go, so to speak - how robust one intends the assertion to be. It is important to realise this, because it means that any *general* statement of the conditions under which one is right to say or think an epistemic modal will have to be quite schematic - or, at any rate, will require just as much contextual determination as the epistemic modal does. Unless one makes reference, in the statement of conditions, to something like 'the

I am on the brink of an error, or a myth, here: the myth being that, whenever we say or think an epistemic modal, we have definite intentions with respect to the level of robustness of our claim. One might say 'It could be that p', and one may as it were think that one intends little or no robustness, i.e. that one is not going out on a limb much at all, but it might turn out that one has just been guilty of a surprisingly elementary confusion, and that it couldn't be that p. In such a case we would say we weren't right to say/think that it could be that p, even if we didn't in

This may make it look like one can never protect one's self from *that* sort of refutation, one must always intend at least robustness with respect to elementary confusions. But this would be dogmatic. There are cases where we become admittedly totally confused and say 'It could be, I don't know, I'm confused', or 'maybe' ('maybe' seems particularly apt here), perhaps in a slightly desperate, or perhaps mock-desperate, tone of voice. This is a characteristic sort of use of epistemic modals, when in a muddle. One shakes one's head, as it were, and stands back from the controls, from judgement. A sort of flailing 'maybe'. Someone might suggest something. "Maybe - I have no idea".

Curious; this sounds perfectly good: 'It isn't so, but could have been'. This sounds queer: 'It couldn't be so, but could have been'.

'Could he have lost?' ambiguous. 'Could it be that he lost?', 'Could he have lost instead?' - and of course this latter can mean various things: epistemic, absence of certain potential defeaters, conceptual-metaphysical.

However, this sounds OK: A: 'Could it be?', B: 'No,

* * *

It is tempting to think of there being the unrestricted metaphysical or subjunctive modality, and then there being restricted subjunctive modalities on that. But of course this won't do for epistemic modality. And yet epistemic modality (truth for all one knows) is temptingly viewed as some kind of restriction on something: p is epistemically possible for agent A iff the conjunction of A's knowledge with p is possible in some unrestricted sense. But what is this unrestricted sense? This is the problem of epistemic space. (Chalmers, 'The Nature of Epistemic Space'.)

What is wrong with this proposal?: p is epistemically possible for S iff S doesn't know p or $\sim p$ (and perhaps

we might add a clause to the effect that p must be within A's conceptual purview).

Obviously, we can't say that what could *really* be the case is not that which could be true for all an ideal knower knows. For then what could really be the case would be one and the same with what *is* the case. (Spinozism again.)

* * *

In Edgington's 'Two Kinds of Possibility', she talks as if the two kinds of modality were:

- (1) Metaphysical necessity,
- (2) Knowability a priori.

Or

- (1) Metaphysical possibility,
- (2) Non-eliminability a priori.

But neither of those captures what we mean when we say:

'The answer to this multiplication is possibly larger than 1000.'

Here one might think that this can be done with non-eliminability a priori if we relax the ideal.

With the epistemic ideal in place, all non-eliminables are metaphysically possible. As soon as we relax the ideal, we get metaphysically impossible non-eliminables in addition.

We might say that we are restricting the ideal, but this does not give us a restriction on a 'modal space'. On the contrary, it expands the space. (More possibilities, less necessities - or rather, more non-eliminables, less eliminables.) *Another* thing we do in ordinary epistemic modal talk, another kind of restriction, is to assume lots of things we believe, thus ruling out their negations. This narrows the space.

Think of knowability a priori as like tautologousness in logic: $\vdash P$. Two things we can do, two forms of restriction: (i) weaken the consequence relation - this corresponds to relaxing the cognitive ideal - and (ii) assume premises (thus talking about $A \vdash P$ instead of $\vdash P$, or equivalently, adding axioms; to extend the analogy further, premises could be contingencies, axioms and consequences thereof could be necessities).

When we judge that something could be so, in some epistemic sense, what is it which keeps us from knowing whether it *is* so or not? (What is the salient knowledge-defeater?) This can vary. In some cases, it can be understood as - roughly speaking - a lack of a priori acumen. In other cases, one is in a position to come to the currently absent knowledge by thinking about things and forming a judgement, though not by a process which we would call deductive or a priori. In other cases, one needs something from outside one's self in order to be able to get at the relevant knowledge. It is tempting to think that this last sort of case only happens with what we call empirical matters, but there are puzzling cases in mathematics and logic which look like this also (and this is basically identical with the problem of apparently Twin-earthable yet non-empirical concepts, which I haven't written much about yet). EDIT

Realizing that there are these quite broadly different types of situations where we use non-idealized epistemic modal statements can help us avoid all kinds of puzzles. For example:

A: It could be that p, for all I know.

B: No, because <some deductive reasoning from A's knowledge>.

A: Oh, you're right. It couldn't be that p.

If one is hung up on the first kind of knowledge-defeater - lack of a priori acumen - such an exchange could seem very puzzling. It might seem like A could have stood their ground here and insisted that their first utterance was correct at the time, since they had not yet engaged in the reasoning given by B. Of course,

that might be ridiculous from a practical point of view. And this might suggest that A was in effect claiming that some further knowledge-defeater was present besides lack of a priori acumen, perhaps something which could only be remedied by empirical investigation. That is how B would be able to refute A.

There are several kinds of possible disagreement, or contradiction, in conversations involving epistemic modal statements. The exchange above is one sort of case. Here is a sort of case where the difference between 'Could be' and 'Could be for all I know' is crucial:

A: It could be that p.

B: No, because q, and therefore $\sim p$.

A: Oh, I didn't realize that q. OK, so it couldn't be that p.

Here, I am imagining that B is really injecting - from A's point of view - new knowledge. Suppose A takes this right on board, trusting B's testimony (for good reason, let us suppose). Then, A must acquiesce. This shows that A's first statement cannot be analysed as 'It could be that p, for all I know'. Rather, if one wanted to put in such a parameter, it would have to be something like 'It could be that p, for all we know between us.' But further developments in the conversation - someone else coming onto the scene, for example - would make this wrong also.

If one ignores this multiplicity of use, one might also become confused generally about the fallibility or defeasibility of epistemic modal statements. And, from the point of view of the disambiguations discussed above - which defeaters to be considered as relevant, and which knowledge base is being discussed (personal vs. communal) - it seems we can identify a class where it indeed makes no sense to talk of someone being sincerely in error. (Namely, the class of uses where the knowledge base is just the speaker's, and all possible defeaters are taken as relevant.)

* * *

Some epistemic possibility statements allude to a structure which is not currently (subjunctively, metaphysically) endorsed. Others, to a configuration of the structure. (I was thinking mainly of finely individuated structures here, so that changing on an empirical identity changes structure, but one can get other versions of this distinction by changing the method of individuation.)

The contextual/relative issue comes up both with structural and configurational epistemic possibilities.

The puzzle with epistemic possibility is its subjectivity, the puzzle with metaphysical possibility is its objectivity.

* * *

The hard thing is to take all these considerations properly on board - to thoroughly incorporate them into one's idea of what a proposition can be.

We instinctively downplay the differences we notice here, always trying to see them as variations on a single theme. And of course there are natural reasons for this.

* * *

Epistemic modals are like indexicals from the point of view of rightness to think.

But like definite, invariable statements (in MacColl's sense).

I am in X. (Time passes.) I am not in X. I was right to think I was, and I was right in thinking I was.

X could be. (Time passes.) X couldn't be. I was right to think it could be - it could have been - but not right in thinking it.

What's similar?

What's different?

In the 'I' case, both believers

Here is a puzzle: it looks like, when we evaluate for rightness in thinking, we use our best lights.

But still, we can't thereby become sure that we are right in thinking what we think. Doing our best is not enough - after all, all that gets our epistemic inferiors is a 'licence to think', not rightness in thinking.

Am I right in thinking this?

What is one asking? Well, for a start, one needs to know if one is right to think it. But there is more to it.

How much more? To be right in thinking Cp , need it be that p ? To be right in thinking $\sim Cp$, need it be that $\sim p$?

Or is it enough that the TO-rightness of what I think is robust with respect to some amount of epistemic improvement in some direction?

Think of the possible 'Ultimate verifiers' of ' Cp ', short of ' p '.

One form of ultimate verification: knowing that one won't (or even can't) know.

This corresponds to knowing that the TO-rightness of one's thought is robust with respect to all forthcoming (or even all possibly forthcoming) epistemic improvements.

* * *

Three modes of rightness and wrongness with epistemic modals:

1. Right/wrong given the situation.
2. Robustness/failure of robustness (to some roughly contextually determined extent)
3. Ultimately right/wrong - removing the diamond yields a truth/falsity.

The difference between (1 and 2) and 3 is clear, but the difference between 1 and 2 is more subtle.

One might think: how can one ever be wrong in the sense of 1? For if one is wrong, one is wrong, so one's situation is simply not up to scratch. Given that situation, there is - ex hypothesi - no further wrongness to be convicted of.

The mistake here is that epistemic situations are being individuated far too finely. One can be wrong to think an epistemic modal given one's situation if one can stop being wrong without thereby changing the identity of one's situation. (The 'can' here is grammatical.)

'Could it be?' 'I don't know' - one reason for answering this way, rather than with 'yes', is that the questioner may clearly want something more robust than you can give them.

* * *

The extent to which the problems of epistemic modals and of objective probability coincide has yet to become clear.

It is probable that p . It could be that p . $\sim p$ falsifies the second, but not the first.

But not so fast! I said that because I have been influenced by the philosophical literature on probability statements.

'There is some probability that p, but p is certainly not the case.'

Just as with epistemic modal language, we have an indicative-subjunctive distinction, a live vs. 'timed out' distinction, at play here.

A denial of an epistemic modal is factive, so to speak, in a way that the corresponding assertion is not.

It could be that p. Compare: The best we can do now is be open to p.

Going from believing that something could be so, to believing that it could not be so, means acquiring the belief that it is not so.

Going from believing that something couldn't be so, to believing that it could be so, means relinquishing the belief that it is not so.

The following principle seems to be - in some sense - at work: It is worse to believe something false than to fail to believe something true.

The funny thing is, epistemic modal language gives us a way of dressing up the failure to believe one way or the other as another kind of belief: that it could be so. The thing which has to be seen is that this extends, and makes more heterogeneous, the concept of belief. (But this is unclear. For when did this extension happen?

Another problem with the above remark is that sometimes we say we don't know whether or not it could be that p. I am talking about cases where we so to speak believe that we are in no position to believe one way or the other. (And this is a very flexible concept.)

It is clear that such an attitude has an important role in life.

To say to someone 'You shouldn't believe that p is not so' does not commit one to 'p'.

To say to someone 'You should believe that p is not so', in the sense which I am thinking of, would seem to commit one to the negation of whatever fills in the blank. (What is the sense I am thinking of? Something like an unconditional sense. I.e. not the sort of case where one says that someone should believe something because they believe something else. But in another sense, it's not unconditional, for it is an open question whether one wants to believe the truth. The sense I am thinking of is that of 'Believe it, because it's true'.)

It is natural to compare 'You shouldn't believe that p is not so' with 'It could be that p', and 'You should believe that p is not so' with 'It couldn't be'. But here we get the puzzle: why 'couldn't be', over 'isn't'?

Often when one is looking at something isn't the case, one uses 'couldn't be': 'It isn't so. It couldn't be, because ...'.

If one had instead said 'It isn't so, because ...', then the blank should really be filled in with (an expression of) the cause of 'p's not being the case. Whereas with 'It couldn't be, because ...' one is talking about reasons, or causes for *thinking* one thing rather than another.

One could put such things after 'isn't' too, I suppose, and people do it quite often, but having a separate form on hand here obviously allows for more fine-grained, less ambiguous thought and communication.

This isn't simply a reasons vs. causes distinction.

p isn't the case, for the reason that ...

p couldn't be the case, for the reason that ...

The first form makes one expect a justification for the fact itself that p is not the case. (He isn't there. The reason being that he had to do X.)

With the second form, such a justification fits as well, but so too would a justification for *believing* that it isn't the case. (He couldn't be there. The reason being that I saw him doing X.) In this latter case, we would call it an epistemic modal. Not the former. If he *didn't* come for some good reason, then he *couldn't* come in the former sense. If there's good reason for *thinking* he didn't come, then he *couldn't* have gone in the second sense. (Note the 'have gone' instead of 'come' - here we see another signalling device in action; the past tense picks up this odd job sometimes.)

* * *

It is though a positive epistemic modal comes with a contextual robustness parameter which is such that: once the relevant level of enquiry is complete, if p is still not refuted, the possibility-claim is weakly verified.

If a possibility claim is weakly verified, that means it cannot be strongly refuted. But it can still be weakly refuted.

If a possibility-claim is strongly verified, it cannot be weakly refuted

Weak verification: verification that one is right to think.

Strong verification: verification that one is right in thinking.

Weak refutation: one is wrong in thinking.

Strong refutation: one is wrong to think.

Confusing terminology.

One can show you were wrong in thinking, but not that you were wrong to think, that it could be that p.) However, if p is refuted before the relevant level of enquiry is complete, then one's utterance is strongly refuted. Not only were you wrong in thinking as you did, you were wrong *to* think thus.)

This remark gives a lot of the 'structure' I need.

Does this mean the following question is systematically ambiguous?: What does one commit one's self to in asserting that it could be that p?

For in one sense, one only commits one's self to it not turning out that $\sim p$ within the contextually determined course of enquiry (though of course, in a sense, this need *not* be *determined* - which is to say, it may be left indeterminate - but there is still something there which is given contextually, so to speak, even if it's not determinate). In another sense however, one plainly commits one's self to more than that: one commits one's self to it being that it could be that p. For if one revises one's view on this, one revises one's view on *this*.

What makes the difference is whether or not we use the concept of 'commitment' to cover epistemic modals themselves (i.e. epistemic modals taken as a *kind* of proposition - which is to say, not epistemic modals as analysed into regular propositions). If we do not, the closest we can get via analysis is the commitment to it not turning out that not-p within some indeterminate but to some extent contextually-indicated course of inquiry. The whole point of this analysis is that it comes in the form of a proposition which is *not* an epistemic modal - that it behaves more like we tend to expect propositions to behave. But with *this* sort of proposition, we can no longer play the game of "weak refutation", or of "changing one's view" about something in some sense stronger than that in which one changes one's view about where one is as one's location changes (for example).

(I do not mean to imply that the robustness parameter need be thought of as a simple quantitative 'amount of inquiry'. It often relates to particular inquiries, and specifically not to others, etc.)

* * *

Earlier I wrote that the puzzle about epistemic modals is their subjectivity, whereas with metaphysical modals, it is their objectivity. But that only gives half the story in each case. With respect to epistemic modals, what I would prefer to say now is:

When we assimilate epistemic modals to subjective propositions (e.g. 'I am in such and such a location'), we are puzzled by their objectivity, by their pointing beyond one's "location". When we assimilate epistemic modals to objective propositions on the other hand, we are puzzled by their subjectivity.

The solution, I believe, is to transmute these assimilations into *comparisons*, and to allow epistemic modals to be themselves.

But is that a task which one can attempt? It seems rather that, once one sees that my solution is desirable, one has thereby taken the decisive step.

What remains, however, is to give the description which supports and fills out such a treatment. That is, the sort of description that encourages one to attend omnivorously to the subjective and objective aspects, so to speak, rather than a description which causes one to fixate on one or the other (or on one and then on the other).

(This sort of situation, where one has to steer between two views by giving up something common to them and forging a new (purpose-built) view, occurs all throughout philosophy. Examples:

- Identity as a relation between objects vs. identity as a relation between names
- Existence as a predicate of objects vs. existence as a predicate of names (modern philosophy has done a little better on this score, it seems, than on identity).
- Analyticity about ("metaphysical", to use a term anathema to the proponent of this view) modality vs. substantive-truth-ism as espoused by contemporary metaphysicians.
- As Anscombe put it, 'the stupidity of empiricism' vs. the wildness (actually, she said something else here) of idealism.
- Cognitivism vs. non-cognitivism in metaethics
- The view that names have senses given by definite descriptions or clusters thereof vs. the view that names 'have no connotation', nothing to their meaning except their referent.
- The view that conditionals can be given a non-circular truth-conditional analysis vs. the view that conditionals are not truth-apt.

Quasi-realism and related views are interesting in this connection (for example, in the case of cognitivism vs. noncognitivism in metaethics). It is as though they go some way toward the sort of treatment I advocate, but then fail in some more subtle way.

Quasi-realism makes it look like, although it is correct in a kind of easygoing sense to speak of truth and the like in quasi-real (what Blackburn calls 'nondescriptive') contexts, it is not strictly correct.

But how unclear! For how isn't that just a labelling issue? Blackburn wants to have more restricted concepts of truth and of proposition and of description, it would seem, than ordinary language gives. And he is welcome to them. But the funny thing is, I possess (and want to possess) very similar restricted concepts myself, but I feel like Blackburn's quasi-realism is a bad view.

* * *

Apriori-aposteriori

Once we become clear about the contingent apriori and the necessary aposteriori, and the connection with "semantic externalism" (Twin Earth, etc.), we have what we need to make better sense of the notion of apriori-aposteriori itself. Positive, "unifying" characterizations tend to have a dark, archaic quality. All that can be avoided. The notion which results is in some sense more structured, more disjunctive, than traditionally thought. Also, the idea that here we have an epistemological distinction, whereas modality is metaphysical and analyticity semantic, has to be treated with care - there is nothing definitely wrong in it, but it tends to create the impression that the apriori-aposteriori distinction is about something **completely** different from metaphysical modality and semantic concepts. (It could be said that, while one is epistemological, one metaphysical, another semantic, **all** are logical concepts.)

The apriori = (the metaphysical necessities minus Twin-Earthable cases) plus thoughts whose negations are, in the host system, essentially counterfactual (e.g. 'This room is as big as it is', 'S is 1m long at t', where S is the standard metre).

For this to work, it is vital to be clear that, when talking of 'Twin-Earthable cases', we are speaking of Putnam-underdetermination, not Wittgenstein-underdetermination. That is, the underdetermination of reference by the language system, not the underdetermination of the language system by **local** neural events, marks and noises, behaviour etc. This is important, because there **are** Twin-Earthish cases in the realm of the apriori, but these are cases of Wittgenstein-underdetermination. (For example, mathematics students: the system they are in now depends on what they are learning, what the teacher will teach next, etc. Their own understanding might fit both with the system they will go on to learn (and are thus already learning) and a narrower one which goes less beyond the fragments they already understand - a simpler

calculus.

To take a concrete example: a logic teacher might start teaching students first order logic with identity, but without explaining what this is and without introducing identity for several lessons. Then the students would actually have a partial understanding of FOL=, although what they understand might fit equally well with FOL without identity. (This is just one sort of case, of course; the "unintended interpretation" need not always be simpler. But this case is particularly clear.)

* * *

The experience of checking a sum is not an empirical verification. There is an important difference between empirical confirmation proper, and a mere case of experience leading to credence.

Synthetic apriori and synthetic aposteriori - what's the difference? Wittgenstein's material in *RFM* about the notion of mathematical proof - having to be easily reproducible, inducing 'spontaneous' concept formation, etc. - is extremely suggestive here. But the discovery of empirical identities, for example, could also be said to involve spontaneous concept formation. (The synthetic apriori as having to do with the form of experience, the aposteriori with its content. This hangs together with internalism and externalism about thought, propositions - however, some aposteriori propositions appear to be semantically neutral, i.e. non-Twin-Earthable.

* * *

There is something very strange about Chalmers's pronouncement that if an agent S knew nothing, all scenarios in epistemic space would be epistemically possible for S; if I've never heard of or met Elton John, is it epistemically possible for me that Elton John exists? Or that he will release a new album this year? A singer called 'Elton', certainly - but Elton himself?

* * *

Notice the nonsensicality of Why-questions about many non-empirical propositions: 'Why are all bachelors married?', 'Why is $2 + 2$ equal to 4?'. It is plainly not advisable to simply reply 'Because we use such-and-such linguistic conventions.' But on Ayer's view, it's hard to see why that is not a sensible Q&A. Whereas anyone can see that the questions are not in good order. The answer Ayer gives might help in many cases (for instance if a child is confused), but it is not really a direct answer. It is used to jolt the questioner out of their question. To make it clear that the matter is not open for question in *that* way.

* * *

It's odd: the apriori/empirical distinction should still be fine, since all that has happened is that we now recognize some necessities which we call a posteriori. But then that makes us wonder: what's so special about the a priori? See, before it looked like we could understand this in terms of some sort of conceptual invariance, but now that no longer seems like a unique marker. The thing I keep wanting to say is: well, OK, experience is pretty much always needed to get anywhere with concepts, although with the apriori cases it's only the *form* of the experience which matters - which is relevant to determining the object of the concept. With Twin-Earthable concepts, however, this does not hold.

Someone may meet someone, then a twin, and take them to be the same person - even though they look, let us suppose, slightly different (so yo can tell, once you learn). This is a perfect setup for a case of the necessary a posteriori. Now compare:

Someone is dancing in a certain way which you haven't seen before, and you form a concept of it. ('That way of dancing.') But then suppose you think you re-identify it numerous times, when in fact you're wrong half the time - there is a distinct but superficially similar way of dancing, which you have not distinguished.

Now, why are these cases so very different?

My first thought: two people can be qualitatively identical, intrinsically. Their distinctness doesn't lie in any

subtlety of that sort. We say there are two not because we can discriminate, but simply because we can put them together and count them. (Note the essential invocation of an intrinsic-extrinsic property distinction.)

There is no such thing as a 'merely numerical' error of identity in the case of universals.

I think it will be helpful to talk for a while as if universals are just like individuals, except multiply located.

Now. Is this possible?: I encounter a universal U and form a concept C of it. But I choke somehow, and the concept I form would actually function better as a concept of another universal V?

* * *

One gloss on the apriori is this: the apriori is what can be worked out by rational reflection - without empirical investigation - given certain concepts, but it is granted that experience is required to acquire these concepts. But what happens to this idea in light of the Kripke-Putnam lesson that our concepts may be internally adequate, and yet inadequate to their objects?

* * *

Miscellaneous

'Modal knowledge' - the very expression is shot through with a certain way of looking at things.

(Nobody uses that expression outside philosophy, and it's *not* because they don't ever refer to what it denotes.)

'Access to modal knowledge' - 'A window to the world of possibility'.

* * *

One thing which must be avoided is Kripke's 'What we *really* mean is ...' move with respect to the subjunctive epistemic possibilities of the table being made of ice, Hesperus not being Phosphorus, etc. ('It could have turned out that ...'.)

Whatever we really mean, we have really said it. (*That* way of using language must not be neglected. It cannot be swept away by a spring cleaning of the language.)

* * *

'Actualizing' in the philosophy of modality and modal logic.

The notion that for all actual truths P, we can get a necessary truth by saying 'P is true at @', where '@' rigidly designates the actual world.

A proper name for the world!

And it's really not clear that it works. Because couldn't *this very world* have been different?

The notion of rigid designation breaks down here.

""@' designates this world at all worlds where this world exists" is not trivially true - it makes no sense.

We are in a muddle. Shades of Liebzniz and Spinoza - the use of '@' as it were invokes the notion of, at limit, an individual concept perfectly adequate to the actual world and no other.

" 'Cows are all white' is false, but not necessarily false. 'Cows are all white at @' is necessarily false."

On one way of looking at things, that just seems wrong. Just as this very man could have done something else, so too could this very world have been such that cows are all white at it.

But suppose someone says: 'No, you only think that because of your inadequate concept of this very world. Worlds have all their properties essentially.' - Well, yes; one can individuate world concepts *that* way too.

We are in the grip of a hellish oscillation between two ways of thinking of non-actual worlds (scenarios) - as a space of things with their own identities vs. a space of shadows with reality in the centre.

(The picture of gaining knowledge as narrowing down one's location, moving toward knowledge of exactly which world is actual - which would be omniscience.)

* * *

'The Sapir-Whorf hypothesis' - as though there were a clear proposal, and we then had to do empirical investigation to see if it were right!

* * *

On the utility of modal judgement

Recent philosophy of modality has come to concern itself with the problem of 'the utility of de re modal judgement'. I get this phrase from John Divers' paper on 'Quinean Skepticism about Modality after Kripke'. Divers takes it that the main thing to be skeptical about regarding de re modality these days is that we still have no good account of the utility of judgements of de re modality (i.e. how things could and must have been, essence).

(Perhaps connected with this is van Inwagen's stuff about there being regular, everyday de re modality, and some kind of puffed up, discreditable philosophical kind.)

The problem with Divers' problem, roughly, strikes me as this: Divers and others talk as though de re modal judgement is a kind of free-standing, or at any rate quite singular, feature of our thought. This comes out in the fact that they look for its utility in a very focused and direct way - as though it has some special, quite particular purpose. But this seems to me a wrong-headed way of proceeding. Think of our judgements as to how things might turn out in the future. As to the utility of these, there is no puzzle there - of course it's useful. To do that, we need a conceptual scheme with moving parts, and we will obviously want to distinguish some propositions as being fixed one way or the other, etc. And what do we need for 'subjunctive', 'metaphysical', modality? A conceptual scheme with moving parts, where some propositions are distinguished as being fixed one way or the other.

(When learning to think in evolutionary terms, one of the important lessons is that not every trait properly explained by natural selection actually has - in itself - any *particular* survival value. And here the notion of a mere side-effect or epiphenomenon is often inappropriate as well.)

* * *

The crudest conceptual analyses in philosophy are often the most enlightening - in metaethics, for example. It would certainly be wrong to think that these crude ideas are only incidentally crude and false. Rather, their crudity and literal falsity is essential to their being illuminating. We are noticing an important analogy - and for an analogy, you need two different things, and so the analysis is bound to be crude and false *qua* analysis.

* * *

Ethical and aesthetic judgements can seem to have a kind of necessary character.

If we "believe" these things, we're typically quite unprepared to think they could have been otherwise while the object of the judgement remain the same in all other respects. And yet they in no sense follow from a "factual" description of the object or situation in question - or so it would seem.

'Imagine this butterfly exactly as it is, but ugly instead of beautiful'. (Phil. Grammar p175 s127.) Despite what I just said, one can actually do something which might be called trying to do this. When I do it, I vaguely imagine sinister connections on the part of the butterfly - though not in any kind of ridiculous literal way.

Suppose we only ever saw butterflies when we were in the most hellish straits, and never saw anything visually like a butterfly in ordinary life - no bright colours, no symmetrical patterns, etc. In that case, would we see that butterflies are beautiful, despite the horrible associations they would no doubt call up? And if not, would this mean we were blind to something? Then shouldn't we also say that we are usually equally blind to the potential horribleness of butterflies? - There is something right in that; think of the way extremely mundane things are often used in horror movies and the like. Chirpy radio ads, for example, during a murder. Freud's essay on The Uncanny (Unheimlich) deals with this sort of thing.

* * *

'What is not the case' as a separate, larger realm from what is the case. But how could that be, since after all it will be the case that these things are not the case?! Well, for the purposes of saying how things are, negative propositions are mostly inefficient and indirect - but perhaps this is badly put: for it seems that this is the very thing which makes them negative propositions.

* * *

I have not paid sufficient attention to the importance of *where a concept is stationed* in the wider system, for the question of its adequacy.

This comes out as follows: I have been concentrating on cases of the necessary a posteriori and have come to the view that these are, roughly, adequate-concept-sustained truths the requisite concepts for which are empirically gotten.

The failure of adequacy in the relevant sense first made itself known to me in such cases, where the failure is connected with bad empirical evidence, lack of empirical evidence, and such things - in short, where the failure was an empirical matter, to be empirically remedied.

But plainly this is but one form of conceptual (in)adequacy relevant to modality. It was dramatized by the object being 'out there', the term being Twin Earthable, etc. But once we see there that certain empirical concepts can be coherent while inadequate to their objects, it becomes obvious that this applies much more widely. Even in mathematics, for example.

(None of this is unwelcome: this is presumably just the way this sort of inquiry should proceed. For I never would have gotten anywhere if I hadn't begun with problems which had been simplified almost to the point of ridiculousness.)

Now, the present issue is very closely connected with the problems of reference and representation, and 'externalist' developments therein. For whether a concept is adequate to its object naturally depends quite heavily on which object that is! And so here the philosophies of modality and reference must shed light on each other.

It may not be going too far to say that Frege-Russell descriptivism about reference and Ayerian apriorism about modality are two sides of the same coin. It doesn't matter too much which one you attack first, because the other will fall along with it. I will have to try to make this clearer.

* * *

Fine vs. course meaning change

Fine: corresponds to change in the conceptual system but without changing the identity of the concept or proposition in question.

Coarse: corresponds to change in the very concepts or propositions - their identity - tied to certain linguistic forms.

Qualitative vs. numerical change of meaning.

'The term changed its meaning.'

Qualitative - the term had a meaning, and this meaning has altered. (The house has changed, in this and that respect.)

Numerical - the term had one meaning, now it has another. (New house.)

So, we've got qualitative vs. numerical change at the level of concepts and propositions.

* * *

An ordinary contingent statement is a move within a system, a grammatical necessity is an artefact of a system, but what about 'synthetic' necessities? Empirical cases - informative identities, e.g. - seem to be like bridges to a new system. What about the case of arithmetic? That does not obviously apply here - although on second glance, one can see a similarity. Think of the effects of learning the times-tables. One memorizes a bunch of synthetic apriori truths which one already has the means of working out on demand - in effect making them analytic, to speak roughly. This of course has very palpable effects. We think differently when it comes time to solve practical arithmetical problems - are able to use simpler, easier movements (patterns), due to a richer toolbox, so to speak.

Often, in the course of some involved apriori reasoning (I am thinking of logic and mathematics), one works something out and then "black boxes" the working out, using just the result. Lemmas. One may then forget these later - not finding them of any particular interest outside of the original context. One can think such lemmas as temporary concept formations, temporary times-table rows, as it were.

With apriori propositions, one person's "structural artefact" is another person's bridge to a new conception. And another use: teaching language to someone who already has the relevant concepts in their own language.

* * *

I distinguish two 'levels' of movement in our conceptual schemes: configurational and structural.

The old, logical view of necessity focused on the former to the detriment of the latter, and thus necessities came to be seen as not involving movement or needing to move in answer to anything. But consideration of cases of the necessary a posteriori - most clearly, cases involving the identity of empirically known individuals - shows this to be inadequate. For here we can move back and forth rather as if we were dealing with a contingency, and yet we *also* want to privilege one of these positions as the right one, and then concentrate of the possible configurations which arise 'within it'. (This moving back and forth can be thought of as moving from one system to another, from an epistemic modal point of view. But another way is to individuate conceptual systems more course-grainedly and to regard 'Hesperus is Phosphorus' and 'Hesperus as not Phosphorus' as both corresponding to possible configurations.)

A good analogy here would be to a game which permits players to formulate new rules, but within limits. So we get a space of possible rule-sets, and within each one, a space of possible states of play. But during play, the rule-set changes.

* * *

A kind of indicative-subjunctive distinction is often used to mark the distinction between epistemic and metaphysical modality: what could be vs. what could have been. It must be noted, however, that both "moods" can be used both ways.

There seems to be a 'restricted', epistemic-on-both-sides manifestation of the indicative-subjunctive distinction in modal talk, where the distinction seems to be purely about tense.

- What could happen; what is reasonable to expect
- What could have happened; what was reasonable to expect

Here also belong uses which point to the absence of certain defeaters and the like.

Furthermore, we often use indicative grammar when speaking of general non-epistemic (grammatical, metaphysical) possibilities: 'a bachelor can't be married', 'thoughts don't have to be conscious'.

Despite this, it is very natural to employ an indicative-subjunctive switch in order to explicate the difference between epistemic and metaphysical modality: 'We could be wrong about p . That is, $\sim p$ could turn out to be the case. But if p is the case, then it could not have been otherwise.'

Here the question arises: why do we use just this grammatical distinction in just this way?

Indicative looks out, subjunctive looks in.

Indicative is open, subjunctive is closed (all bets are off).

Epistemic possibility statements permit one to envisage various alternative conceptual structures. Metaphysical possibility statements presuppose the "right" structure and concern the available configurations. (This is not as cut and dried a matter as it perhaps sounds.) For the most part, it seems right to say that epistemic modality is more flexible than metaphysical (although of course *both* allow movement), but on the other hand: the contingent apriori, the proposition that things are different from how they actually are - essentially counterfactual thoughts. So in *that* respect metaphysical modality is the more flexible. But this sounds like nonsense - as though the important difference is being completely skirted over..

With what *is* the case, all is fixed. With what *could have been* the case (in the metaphysical sense), we hold the conceptual system fixed, but not the configuration. What *could be* the case is something else entirely.

* * *

There are ordinary epistemic possibilities which are not ideally conceivable (or: which can be ruled out apriori), and there are ideal conceivabilities which are not metaphysically possible.

Here we get a kind of Russian doll picture, with metaphysical possibility as the smallest.

But the contingent apriori makes trouble for this: it is metaphysically possible that the metre stick is less than 1m. Well, that sounds funny, in the indicative. ('Metaphysically possible' is a highly technical term - and it's not just a label but a grammatical construction, so to speak.)

This is actually, surprisingly, closely related to the "Elton John problem". (For it to be epistemically open to someone that Elton John released a new album, they need to be able to think of Elton John, it would seem.)

* * *

EDIT

What could have been the case, in the sense of metaphysical possibility, pertains to the possible configurations of "the right system" (both a priori certified and empirically adequate). And this includes essentially counterfactual configurations, e.g. S could have been shorter than it actually is, Julius (<- descriptive name defined by Evans) might not have invented the zip.

We can then talk about all the possible configurations of all apriori certified systems.

We can also talk about the possible configurations of the "right system" which are not apriori false - and this just is the set of possible configurations of the "right system" minus those which are essentially counterfactual.

The epistemic-metaphysical distinction in modality is unclear because it runs together numerous things.

Kripke's work, and developments in 2D semantics, lead us to compare the space of things which could be true for all that can be known apriori, with the space of what is and could have been the case - the space of metaphysical possibility.

In moving from one space to the other, two things are happening simultaneously. One is moving from indicative (consideration as actual) to indicative plus subjunctive (actual + counterfactual). One is also introducing maximal empirical knowledge, in addition to maximal apriori knowledge. The first thing adds to the space (hence the contingent apriori - apriori impossibilities which are metaphysically possible) while the second subtracts from it (hence the necessary aposteriori - apriori possibilities which are metaphysically impossible).

So what about the 'Russian doll' picture of modality? Here is one way in which it can legitimately be applied:

A space of thinkables - thus including things which an ideal reasoner could rule out apriori, but which a less ideal agent might not be in a position to rule out, or even just things which we haven't worked out yet (the answer to a sum - it might be even, might be odd, let's see). Then a smaller space of things which are epistemically possible for all one knows apriori. Then a smaller space of that which could be the case for all one knows. Then 'what could be the case for all which can be known' is just what *is* the case (unless one has a philosophy which says that there are truths which are in principle unknowable).

In this case, counterfactuality, subjunctiveness, is never introduced. Now, perhaps there is a separate application of the Russian doll picture if we talk about all possible configurations (what is plus what could have been).

Let us start with the smallest doll and work out. The smallest doll is the space of metaphysical possibilities - the space of configurations of "the right system". To get to the next largest doll, get rid of empirical modal knowledge - the necessary a posteriori. Then you get a space of apriori possible systems, and all their possible configurations - so you get things like 'Hesperus is not Phosphorus', 'Water is not H₂O'.

Needless to say, both these doll-picture applications can be inverted by switching from possibility to necessity.

* * *

We can use our system to describe a world in which we would do better to have a different system.

This corresponds in some way to when Chalmers talks about 1-intensions - if all the clear, drinkable, watery stuff were XYZ, etc. But isn't this only possible due to a kind of overlap? Doesn't such talk only get us *so far* away from our current system? Here the picture of possible worlds on the one hand, and descriptions (representations) of them on the other, is apt to play havoc with us.

At bottom, it's all about the actual world in my view. The actual world is the paradigm of worldhood. And not only that!

Am I saying that the multiplicity of possibility does not come from the world, but from its representation? No, I wouldn't say that - for that ignores the connection between the world and its representation. World -> representation -> the multiplicity of possibility. In any case, such "high level" talk really does not lead very far - is not very important for what I'm trying to do. What needs attention is the more concrete stuff which makes one want to say it.

Counterfactual modality concerns range of movement. Subjunctive metaphysical modal statements delineate range of movement. Live epistemic possibility statements envision structures and configurations, as it were for potential use. But here we also get a kind of range of movement - a wider one. Everything at all conceivable. We might say that ideal conceivability - the apriori - cuts that down in one way, and the empirical - empirical concept-formation - cuts it down in another (and these 'ways' themselves will admit of subdivision). That can all be understood as applying to epistemic modality in some sense. (Moving to metaphysical modality changes the game again - one gets configurations which are essentially counterfactual. The contingent apriori.)

It is confusing because there are multiple important distinctions in play. There's the all important distinction between subjective and objective modality, or epistemic possibility vs. metaphysical possibility. But: this very difference is the product of two factors which we want to separate: the empirical and the apriori.

Take bare, totally unideal epistemic possibility with minimal cognitive resources. Think of empirical knowledge and reasoning power as two resources which can now be added. Adding them yields both an improved base for epistemic modal judgement, as well as a richer endorsed conceptual system whose configurational possibilities provide a base for metaphysical modal talk.

Perhaps the greatest danger here is in conflating the apriori-empirical distinction somehow with the epistemic-metaphysical. (Something like this seems to happen in Edgington's paper.)

Modal rationalism vs Modal empiricism, following from

The appearance of two sides here comes from unclarity.

The rationalist says that an ideal reasoner can all by themselves tell possible from impossible.

The empiricist says that to tell possibility from impossibility, one often needs empirical knowledge.

But what if, to be an ideal reasoner, one needs empirical knowledge?

Consider the modal conception of validity in this connection, and what this becomes in light of the necessary a posteriori.

I don't mean that the notion of an ideal reasoner is clear - but think of the good reasoners we know. They couldn't be any good without empirical knowledge.

But I'm not trying to say something causal - that as a matter of fact, it's practically impossible to get any good at reasoning without having empirical knowledge. (That wouldn't have this kind of interest.) Rather, it's that if you're an ideal reasoner you *must* have some empirical knowledge. The hypothesis of ideal reasoning without empirical knowledge is a mere notional fantasy, which can't be worked out in concrete terms at all.

Here one might think the ability of computers to reason much better and faster than us about things like arithmetic, as a potential mark against what I'm saying here. But think what a computer would have to have on board in order to reason well about empirical matters. There is in fact a growing discipline of designing 'ontologies' for computational application - usually heirarchical structures for representing information. (Also, the much-hyped 'semantic web'.) Consider what these ontologies might look like in special branches of biology, for example.

From 'Lois Lane and Clark Kent were in the same room' we may deduce 'Lois Lane and Superman were in the same room'.

One does not need the identity as a premise. This of course does not tally with formal logic, but it is true according to the modal conception of validity, if the modality is taken as metaphysical. Furthermore, it

sounds intuitively right. Sure, one may say that here one needs background knowledge that Clark Kent is Superman. Of course. But this background knowledge is of a certain kind.

Compare the statement that from 'All teachers drive' we may deduce 'N.N. drives' - although of course we need the background knowledge that N.N. is a teacher. That sounds wrong. This seems to have something to do with its being a contingent fact about N.N. that he is a teacher. Furthermore, it is contingent that there is any N.N. to speak of. Thus it is not the case that it is impossible for 'All teachers drive' to be true while 'N.N. drives' is false, and so the inference is not valid.

Now compare: from 'All men are mortal' we may deduce 'Socrates is mortal' - although of course we need the background knowledge that Socrates is a man. That sounds wrong too, but in this case it's not clear that we can say that it's merely contingent that Socrates is a man. Perhaps it is of the essence of Socrates that he is a man. But this, we might want to say, just ensures that 'Socrates is not a man' or 'Socrates is <something incompatible with manhood>' will not be true. It still doesn't ensure that we can say anything which implies the existence of Socrates. He might not have existed. Thus 'All men are mortal' could have been true while 'Socrates is mortal' might fail of reference. So the inference is not valid.

It could be said that it's not strictly accurate to speak of the necessity of identity - since identity statements usually imply existence. Really, it's the impossibility of distinctness of identicals. But of course, identities are for this reason to be distinguished from contingent statements - and it is very easy to slip into calling this distinguishing feature 'necessity'. But really it's the impossibility of being truly negated. (Not metalinguistic negation or that-phrase negation, but internal sentential negation.)

I initially said that identity statements imply existence, but really that is dogmatic. (I added 'usually' subsequently.) Why shouldn't there be a non-committal use of ' $a = b$ ', which is true not iff a and b (exist and) are identical, but iff it is not the case that a and b are distinct. Or we can allow talk of nonexistent things and say that a and b 's being identical doesn't mean they exist. Indeed, the case 'Clark Kent is Superman' is in reality of this kind - although I usually use that example as though the fiction were reality.

Suppose someone actually believes the story of Clark Kent and Superman. Someone might inform them that Superman doesn't exist, and in this case it would be natural for them to conclude that Clark Kent does not exist. The fact that an identity can sustain such an inference shows that identities do not presuppose existence in the way other sorts of statements do, for if they did, the identity would have to be abandoned as soon as Superman's existence is abandoned, before one could infer that Clark Kent's existence is to be abandoned. Compare 'Clark Kent saw Superman'. If we learn that Superman doesn't exist, we have to just drop that. No one really saw Superman. We don't infer that Clark Kent saw something nonexistent, for example.

* * *

One feels that quasi-realism (and similar views) go wrong so to speak at the periphery, in the way they connect the dealt-with area up with the rest of life. So long as you're just talking about logic, quasi-realism is as it were equivalent to permissive realism. But when you start to look at statements like: 'What we really care about is truth', 'Here we are only interested in the most objective features of reality' and other things, it would seem that there is a huge difference between quasi-realism and permissive realism. ('Permissive realism' is the view which agrees with Blackburn's motivating logical points, but which talks of different kinds of truth, instead of quasi-truth.)

GETREF Blackburn had some famous disagreement with another writer about whether language 'has a downtown'.... well... these really aren't things one can profitably argue about, unless one is willing to stop being hung up on that particular idea. To assume that the question of whether language has a downtown is clear enough so that philosophers can debate it and perhaps settle it, is precisely what one must not do. There is *no* basis for such an assumption, logically or historically.

* * *

Getting clear about metaphysical modal statements does two things on the propositional classification front, simultaneously: we mark out propositions according to their metaphysical modal status, and we also mark out statements *of* metaphysical modality from other statements of similar or identical form.

* * *

De Re - De Dicto and related: ICs introduced, bel reps, rigid designation
so this will probably fit in with ref and rep, some of it should be put next to, end of one sec to start of
other, or subsects or something

The way of speaking used in modal dialectic - 'that very thing, *independent of how you describe it*' - is very weird. How does *that* formulation do what's needed? For it is true, on a de dicto reading, to say that, necessarily, the tallest man, independently of how you describe him, is the tallest man. (You don't have to be described any particular way to be the tallest man.)

* * *

'The de re/de dicto distinction' can appear in two lights. De re can appear as a special case of de dicto, via existential generalization: the name 'Clark Kent' in the de dicto proposition 'Lois believes Clark Kent works in an office' has a dual role - in (adapted) mediaeval terminology, it has dual supposition: it supposits for an object, and also for a concept or mode of presentation or something like that. We can bring this out by splitting the job up into two occurrences: 'Lois believes of Clark Kent, via her Clark Kent concept, that he works in an office.' Now, de re cases can be seen as ones where the second part is not specified - as a kind of existential generalization from the de dicto case: 'Lois believes of Clark Kent, via *some* concept, that he works in an office'. And in this case we may of course substitute 'Superman' for 'Clark Kent' *salva veritate*.

On the other hand, in modal contexts, de dicto can seem to be a special case of de re: 'Socrates is necessarily F' is de re, whereas 'It is necessarily true that Socrates is F' is de dicto. But if we turn the operator into a predicate we get: '"Socrates is F" is necessarily true', which has the same logical form as 'Socrates is necessarily F', and thus comes to seem like a special case of de re modality, namely where the res is a proposition or sentence.

However, there is a subtle danger with de re belief reports to think that the main subject term in some sense *purely* contributes an object, acts as a 'mere tag' - this is nonsense. Its conceptual 'content' or 'associate' will appear in order to pick out the object, but will not play the dual role that it might in de dicto belief reports: to pick out the object, *and* to say how the believer picks out the object. (Likewise with reports, e.g., about what propositions or sentences say.)

* * *

Pierre believes, of London, both that it is beautiful and that it is not beautiful (because he splits London).

Pierre believes of London via Concept-Londres that it is beautiful, and believes of the same city via Concept-London that it is not beautiful.

Or if we aren't in a position to specify concepts:

There are two distinct individual concepts X and Y of London such that Pierre believes via X that London is beautiful and via Y that it is not beautiful. (Distinctness of *concept* is what matters here, when judging Pierre's rationality.)

The difference between single and dual supposition in belief reports is that with single, the concept is used for us only - to get to the object as it were - and we just get an existential generalization about the other's conceptual situation. With dual, the concept also is understood to be the one via which the belief is had.

One gets into all sorts of confusions if one "forgets" that for any particular object that we believe in, we will always have a single master concept, even if we have subsidiary aspect-concepts, *unless* we are in a Pierre/Lois Lane-like splitting situation. With this in view, we see how individual concepts can ground necessary a posteriori identities. (Empirical concept formation.)

* * *

t designates rigidly iff it picks out the same object *o* in all possible worlds at which *o* exists. But how do we

specify *o*? With an expression, with a concept.

What's right about the idea that names are 'rigid designators' is that names obviously function differently from definite descriptions, and in this sense have no content - no *general*-conceptual content. However, that doesn't mean that we don't have an individual concept which we attach to that name (though of course this attachment is not a reference-relation).

Concepts are elastic, and individual concepts are no exception (the elasticity need not appear at the level of actual extension). Moving out into counterfactual space, you don't suddenly come to a stop. (I am thinking of, e.g. the question whether I could have been a computer scientist, or a sporting champion, or female, or a neanderthal, or a dog, or a planet... by this stage it's all over, let alone when you get to, say, numbers. However, we can employ structural analogies and say things like, 'If you were a house, you'd be ...'. We might call a very singular person 'prime' in analogy to the mathematical sense, to mean that one can't as it were think of them as a product of factors other than themselves and unity.)

* * *

Names are tied to individual concepts, concepts of objects, so it is immediate that they "designate the same object in all possible worlds". (Here individual concepts are being individuated externally: my concept of water is not the Twin Water concept, since they are of different objects.) Proper names are tied to such concepts, so of course they are "rigid designators" - a proper name, if it designates (i.e. if its associate concept has extension) designates a particular object (that's just grammar) and so of course, unless we equivocate, it designates the same object *no matter what* (that's just what it means not to equivocate with a name). Thus, a fortiori, it designates the same object "in any possible world where that object exists".

* * *

'A believes that *p*' is used in many ways.

For a start, it does not always imply 'A believes the proposition that *p*'. (Kent Bach's puzzle.)

Let us start with the subject-predicate form.

A believes that *s* is *P*.

I begin with a *de re/de dicto* distinction, or perhaps more clearly, Quine's believes-relational and believes-notional. Believes-relational carries ontological commitment, believes-notional need not (except to the belief, or proposition believed).

Now, within believes-relational, for the subject-predicate form, get two possible readings:

(i) A believes, of *s* via concept-*s*, that it is *P*.

(ii) A believes, of *s* via some concept, that it is *P*.

Here it might look like we should include, for completeness:

(iii?) A believes, of some object via concept-*s*, that it is *P*.

But it does not seem that our simple form is ever used to mean this (and no more). When we want to say just this, we use a more complex form. And anyway, if we are individuating individual concepts partly by object, then we will know that concept-*s* is of the object *s*. When I say individual concepts may be individuated *partly* by object, I mean: internally different concepts of the same object may be distinguished, as well as internally similar concepts of different objects - so the identity of the concept, on this conception, involves both its internal nature and its extension. In a word, this identity is a logical product of identity of form and identity of reference.

(Just as there are two permutations and expansions for the subject term, we can do the same thing for the predicate term, yielding a total of $2^2 = 4$ readings of the believes-relational kind. For a two-term relation,

there would be $2^3 = 8$.)

Believes-notional:

A believes (the proposition that s is P)

And now we get the issue of different ways of individuating propositions

- Same concepts, same arrangement
- Concepts of same objects, same arrangement
 - And here there may be different ways of individuating the objects
- Some kind of purely internal equivalence
- Some kind of external equivalence

Now, what is the point of what I am saying here?

It is not to lay out all the possible disambiguations of a simple belief report.

Nor am I concerned to maintain that every time someone makes a serious report, they mean just one of the disambiguations.

The use of the 'of ... via ...' form above shows a way of representing things which is finer grained.

(In language, when you've got simple forms and pictures doing a lot of work, then their working, their application, will be anything but simple. Move to a more articulate form, or picture, and the possibilities of misconstruing the application are diminished.)

* * *

On Method

From W's *Notes for Lectures on "Private Experience" and "Sense Data"*: The philosophical problem is: "What is it that puzzles me in this matter?".

Could I give, as a general aim: a better overview of the use of modal expressions? No. The modal use of expressions? Unclear.

I want to make many meaning-distinctions for free, and that is one reason why I am not doing "semantic theory". (Wittgenstein, Notebooks: 'One of the chief skills of the philosopher is not getting involved in work which does not concern him.')

* * *

When you find yourself wanting to propose a philosophical hypothesis, or thesis, try instead to articulate what made you want to propose it. If this is done effectively, nine times out of ten the thesis will fall away as superfluous, or will become a mere handle or mnemonic. ('Existence is not a predicate', 'Identity is not a relation'.)

* * *

I've recently started to see how much Quine and I are in agreement on something, namely: that the notions of meaning and proposition are questionable material for systematic, scientific theorizing. Only, we draw opposite morals.

Quine's response: so much the worse for 'meaning' talk in philosophy.

My response: so much the worse for systematic, scientific theorizing in philosophy.

* * *

People talk about Kripke having distinguished necessity from a priority. He talks like that himself, and fair enough too. But it sounds hollow to me when people summarize the results of Kripke's work by saying that *apriority* is an epistemic notion, *necessity* a metaphysical one. This has a kind of stitched-up, textbookish blandness to it which seems to me immensely inappropriate, although I admit there is no definite mistake in it. (The closest thing to a mistake one could point to is the unclarity of the signification of 'metaphysical' here.)

There is a temptation to try to get out of all our present difficulties by means of 'subject matter clarifications'.

* * *

A set of philosophical doctrines can be like a plaster cast which sets us in a position such that we are prevented from making the problematic movements. It will soon be noticed that one's movements are restricted - the plaster naturally cracks. One can either refrain from such movements (Quine), or one can then keep rebuilding, allowing more and more room, until one ends up with a weirdly shaped thing which simply does nothing. (A typical day in the office!)

* * *

From Kant's *Critique of Pure Reason*:

For, as a doctrine, that is, as an endeavour to enlarge the sphere of the understanding in regard to pure a priori cognitions, philosophy is worse than useless, since from all the attempts hitherto made, little or no ground has been gained. But, as a critique, in order to guard against the mistakes of the faculty of judgement (*lapsus judicii*) in the employment of the few pure conceptions of the understanding which we possess, although its use is in this case purely negative, philosophy is called upon to apply all its acuteness and penetration.

* * *

Our task is not to ground anything, but rather to identify the ground as the ground. (Wittgenstein.)

* * *

One thing which makes our task difficult is that what at first glance might appear like our guiding questions actually misguide us

'How is reference determined?'

'Is necessity a robust feature of reality, or merely ...' (Realism vs. antirealism.)

'Can modality be reduced, or must we take it as primitive - or eliminate it?' (Paradox of analysis.

Indeterminacy of the predicate '... is a modal notion'.)

* * *

Kripke's most important ideas about modality were not *explanatory* (for example, the notion of rigid designation explains nothing about names, and the causal story is really not to the point or very general). He reminded us that we speak this way, that we judge that way, and in this way showed that earlier explanations were misguided. (And this need not suggest that we need better *explanations*.)

* * *

The central flaw in much of what I have written is that it has the unintended appearance of giving a kind of unexpected grounding for modal discourse. Perhaps it also might look as though what I'm saying might change our evaluative criteria, cast doubt on earlier beliefs about what's possible, etc. All of this has to be avoided. But it isn't satisfactory to simply insist 'No no no, that's not what I'm doing at all' - this has to be

made evident. And the way to do that is to be better (more effective) at what I'm really doing.

Very often, my problem is that I don't know what my problem is. I have to take it easy about that, and just look for frontiers where they arise.

* * *

It is important that I am not offering analyses, nor talking (much) about possible worlds. The advantages of this last thing become most apparent when one considers the problems at the 'Foundations of Two-Dimensional Semantics' (cf. Chalmers' paper by that name, and also 'The Nature of Epistemic Space'.)

Instead of formulations of truth-conditions, and meaning-equivalence claims, I favour formulations like these:

- When one says ____, one is as it were proposing a ____.

- X is like A, whereas Y is like B.

Analogies, connections, dissimilarities, etc.

But none of this means I'm hostile to semantics! I'm just trying to employ a different sort of method, which I feel is required for the problems I most want to make progress on.

Speaking roughly, my view is that with these problems, one has to do something other than semantics or analysis, instead letting all the different kinds of propositions which we learn to distinguish stand irreducibly apart. This is not to say there won't be all manner of very close connections, but the general tendency is to make more distinctions, bringing to bear more and more 'primitive notions' in an attempt to gain a clear aerial view of a previously foggy landscape, etc. It is in this way that my method differs most sharply from those of contemporary semanticists and metaphysicians.

* * *

From Notes from Wittgenstein's Lectures on the Foundations of Mathematics:

One of the greatest difficulties I find in explaining what I mean is this: You are inclined to put our difference in one way, as a difference of *opinion*. But I am not trying to persuade you to change your opinion. I am only trying to recommend a certain sort of investigation. If there is an opinion involved, my only opinion is that this sort of investigation is immensely important, and very much against the grain of some of you. If in these lectures I express any other opinion, I am making a fool of myself. (p103.)

* * *

My work will go better if I don't put too much weight on identifying mistaken views. This impulse is understandable; we want to isolate the philosophical poison, and contain it to certain propositions (forms of words). But this is not always a suitable way to proceed. The poison saturates forms of words which we want to keep using.

* * *

It is to be remembered that the beginnings of modern, formal-logic-influenced semantics came out of the classical conceptions of conceptual analysis, which in turn came out of a desire to use conceptual analysis to solve puzzles and problems. In the present day, however, this aspect is very often left behind. (And that doesn't have to be a bad thing - it might just mean a change in interest, a change in the basis of the activity.) Nowadays it is common to hear talk about semantic theory being used to give scientific explanations of meaning and linguistic behaviour. Sometimes it can even be made to seem as though philosophical puzzles just are things which make trouble for these theories, rather than their reason for existence.